

Iraqis must not use gas again

The welcome ending of the war between Iran and Iraq should not blind governments elsewhere to Iraq's use of poison gas — and to the need that it should not be allowed to happen again.

THE prospect of at least a temporary halt in the eight-year war between Iran and Iraq is welcome not only in the Middle East but also far from the Gulf. The great damage the combatants have inflicted on each other for no great gain on either side is bad enough to contemplate. The ever-present danger, during the past decade, that the conflict would be enlarged has been hair-raising; at different times, one adversary or the other has seemed deliberately bent on doing just that, by hostage-taking and attacks on merchant shipping. Although the conflict has mercifully been contained, non-combatants of all kinds have been sullied by it. Too many states whose wider interests were threatened by the war have emerged as partisan arms suppliers eager to tilt the balance one way or the other, while the discovery that the Reagan White House, with or without authority, was seeking to trade arms for hostages has unhappily renewed the corrosive atmosphere engendered by earlier scandals. But relief that these tensions may now be abated should not blind us to one particularly sinister aspect of the war between Iraq and Iran — the unconvincingly disputed use by Iraq of poison gas against Iranian troops and, possibly, civilian targets during the past few years.

That there is something peculiarly horrible about death by poisoning on a battlefield is not what matters most. In respect of those who die or are injured in battle, it probably makes no difference in the long run whether the damage is done by bullets or by gas. What does matter is that Iraq is a signatory of the Geneva Convention and its protocols which, for practical purposes, formally require that signatories should not use chemical gases as weapons except by way of retaliation in kind.

Iraq's use of gas during the Gulf war is therefore a sign that even states not ranked as major powers can shrug off solemn international obligations when it suits them. That in itself is a chilling precedent that should not be overlooked in the general euphoria that the war in temporarily ended. One Iraqi defence against the pressure that should now be exerted by the major powers against Iraq will no doubt be that nothing was done to penalize Israel for its equally lawless attack on the Tamuz research reactor in 1981 (see *Nature*) 291, 523; 1981. The reply will have to be that, in international relations, one injustice does not sanction others.

The precedent set by Iraq is a poser for the major powers because it will invite imitation by other states precisely when they are trying (at Geneva) to work out among themselves means of turning the Geneva convention as it applies to them into a verifiable ban on the manufacture of chemical weapons. The plain truth is that the technology of the manufacture of poison gas, and the munitions required to carry it against targets, is among the easiest to acquire. States that would not dream of becoming independent of friendly suppliers in, say, battle tanks, can realistically seek to tuck away a clandestine chemical plant making derivatives of phosgene or even nerve gases. On the face of things, Iraq has shown would-be imitators how to do just that. And how to get away with it.

The United Nations, showered with unaccustomed approbation for having organized the truce, has a crucial part to play in the months ahead. Its own unsolicited documentation of Iraq's

transgressions appeared as the truce negotiations reached their most delicate stage, and were no doubt underplayed in the hope of securing Iraq's reluctant signature. Quite apart from the need for diplomatic pressure on Iraq, there is now an urgent need of more detailed information about the means by which Iraq manufactured its chemical weapons and the uses it made of them in the war. The second point is especially important. Iran has understandably protested vigorously about the numbers of people killed and injured by Iraq's poison gas, in the process creating the impression that it was an effective weapon of war. But the chances are that the truth is quite different; for all its horrors, poison gas is more a blunderbuss than a rifle. The threat that it will be used in a war such as that in the Gulf is most likely to encumber combatants with the requisite logistical preparations. Thus an objective evaluation of the uses made of gas during the past eight years may help to persuade other states that this bizarre game is not worth the candle, even in military terms.

But pure reason will not suffice for dealing with Iraq. For that purpose, it is essential that Iraq's international associates, and especially its conventional arms suppliers, should resolve that the use of gas in violation of the Geneva Convention should be penalized effectively. The danger, even when the truce comes into effect, is that arms suppliers will continue calculating how their trade will help shift the tactical balance between Iran and Iraq even though the war is technically over, compromising their long-term interests in the process. Governments elsewhere should recognize that seeming to condone Iraq's use of poison gas by continuing to supply arms of other kinds without demanding a compensatory undertaking to dismantle existing facilities for producing poison gases will be as short-sighted as to seek to trade arms for hostages. □

Nuclear togetherness

Distinctions in the United States between civil and military nuclear power may soon be blurred.

ALTHOUGH the past few years have not been auspicious for reactor-builders in the United States, let alone for those who would develop new designs, the US Department of Energy seems about to embark on such a course (see page 558). The department's plans are born of necessity: new facilities to produce tritium and weapons-grade plutonium are required to replace ageing reactors whose life is being stretched to meet production needs. Two reports from the National Research Council (the research arm of the US academies) show clearly that it is no longer safe to limp along in this way. Many of the serious safety issues the council has identified are unrelated to the age of the military reactors, but it is also inevitable that the risks will be magnified as the military reactors grow old. Current arms control negotiations notwithstanding, something must be done.

While the need for new reactors is military, it is inevitable that the Department of Energy, which also has civil responsibilities, should think of killing two birds with the same stone — winning civil as well as military benefits from the very considerable sum it will have to spend on new military reactors. After all, the



department is the chief inheritor of the old Atomic Energy Commission, partially dismembered in the early 1970s (by the creation of the Nuclear Regulatory Commission) and which happily, perhaps even complacently, combined responsibility for the production of military fissile material with that for the sponsorship of civil nuclear industry. The underpinning doctrine, perhaps more symbolic than real, is important: scrupulously keeping the production of military fissile material in civilian hands, but rigorously separate from civil reactors, has been taken as a demonstration that the old friendly atom and the threatening bomb were separate, but that civilians were in charge of the material of which bombs are made.

But now the Department of Energy is brooding on plans that would blur the fine line between the civilian and military control of nuclear reactors. It is considering using one of its new plutonium and tritium production reactors to help with the development of civil reactor technology and, indeed, in the production of commercial nuclear power. The plans are motivated politically as well as financially. Senator James McClure of Idaho has been a vocal advocate of projects that would find a home at the Department of Energy's test facility in his state.

The Idaho Testing Station already houses the Experimental Breeder Reactor II, a liquid-sodium-cooled reactor with extremely attractive passive safety characteristics whose all-metal fuel can be burned more completely and reprocessed on site, thus offering a promising commercial power design. Its designers are not to blame that their project has been tainted by the failure of the Clinch River Project.

McClure's ambition is that Idaho should also become the home of another promising commercial reactor design, the modular high-temperature gas reactor. He has argued that the United States needs an alternative to the heavy water designs that have been used to produce tritium at Savannah River. He has taken pains to point to possible delays in commissioning a new heavy water reactor. His argument is probably well-founded, if not entirely innocent of special pleading. A second tritium source avoids scheduling problems and issues of vulnerability.

But the high-temperature gas reactor is not nearly as mature a design. Development alone will send the cost of the new reactor skyrocketing, which has discouraged some potential supporters.

The compromise — and a sensible one in some ways — is to help offset the cost of developing the new technology by concurrently developing technology for commercial power generation using the same or similar reactor designs.

Using military development for civilian applications has a long history in the United States. Civilian aviation was propelled by advances in military aircraft design, and certainly the Air Force's Titan 4 rocket is likely to launch civilian satellites, with everyone the better for it: the military gain a little extra money to offset development costs, and commercial or scientific interests get a desired product or service without having to pay the full cost.

The more insidious part of the Department of Energy's new plan is that of using the same reactor to produce electricity as well as tritium. This, supporters will argue, will lower the development cost of the reactor, and make use of valuable heat. But assigning to a power plant such disparate goals has long-term implications. It is the thin end of the wedge, bringing a military application to an essentially civilian function. Even if the safety issues raised by operating a dual-purpose facility can be resolved, the question remains whether joining both operations under one roof would be acceptable. If nothing else, it would be a rallying point for opponents of nuclear power, who have legitimate concerns about its use for electricity generation without adding the inflammatory role of maintaining mass instruments of destruction. Although it is gratuitous to suggest that the next step will be to use commercial aircraft as backup weapons-delivery systems, it is fairly certain that such charges will be heard if the Energy Department perseveres in its plans. □

Nuclear muddle plan

The UK government should delay the sale of its electricity monopoly until its mind is clearer.

THE British government is faced with a different kind of nuclear conundrum, but one of its own making: how to arrange that there should be a nuclear programme of any kind under the intended arrangements for the electricity supply industry, which is to be converted from a public monopoly or nationalized industry, into a clutch of private monopolies under legislation promised for November. From the outset of what has been, in many ways, an imaginative policy for the disposal of publicly owned assets, it has been clear that the dependence of the United Kingdom to the tune of 15 per cent on nuclear power for electricity generation would be a snag. With a few strokes of a very broad brush, the Secretary of State for Energy, Mr Cecil Parkinson, last February described the way in which he planned to carve up an industry whose successes (and failures) have stemmed from its high degree of integration. According to the plan (see *Nature* 331, 466; 1988) there will be two generating companies (provisionally known as 'big G' and 'little G'), the larger of which will be responsible for nuclear generation. There will also be several regional distribution companies, free to generate their own electricity if they choose, who will between them own the national grid, by means of which efficient generating plants can be used preferentially.

These arrangements will apply in England and Wales, with broadly similar schemes in Scotland and Northern Ireland. The government intends to ensure a place for nuclear generation in this new pattern by requiring that the regional distribution companies should buy a certain proportion of their electricity from 'big G', and that a certain proportion of that should be generated from uranium. It will quickly be appreciated that this plan is an awkward compromise between the government's consistent adherence to the doctrine of the marketplace as the only sure arbiter of economic good sense and its admirable belief that Britain (and other similar countries) will need to rely on nuclear generation for the production of a substantial fraction of electricity production.

The most cogent criticism so far has come from the House of Commons Energy Committee, by no means a part of the anti-nuclear establishment. The committee was understandably concerned that coal and nuclear energy were being dealt with unevenly — British Coal (still nationalized) is to lose its protected market, but nuclear generators are to be given one for the first time. It also complained that the arrangements so far described only in outline will make it difficult to know what will be the cost of the government's continuing support for nuclear generation, and to know who is being required to pay for it.

What emerges from the fine print of the committee's inquiry is that these important questions have not yet been thought through. But the practical difficulties of pretending that the electricity supply industry is a series of independent private companies in competition with each other when one will be required to shoulder what may be the extra costs of generating a certain (variable) proportion of its electricity from a specified source makes a nonsense of the notion of competition: 'big G' will be required to operate nuclear plants if it proves to be uneconomic, but presumably there will be nothing to prevent 'little G' from joining in if the economic balance should tilt the other way. Before November, the government may yet be glad to embrace the Energy Committee's suggestion that a more transparent way of entrenching nuclear power would be through a third generating company.

But that is only one way in which British plans for privatizing electricity are half-baked. One has only to note that it had not been (last month) decided how prices were to be regulated to conclude that the best course would be to postpone the whole scheme for selling the industry for at least a year. □

US looking for short cuts to speed drug approval

■ AIDS epidemic forces new approach

■ Critics say changes just cosmetic

Washington

THE US Food and Drug Administration (FDA) is drafting regulations to streamline the review process for drugs to treat life-threatening illnesses. The FDA began the process of paring down the regulatory procedures for drugs to treat diseases such as AIDS, heart disease and some cancers last week, prodded by the Presidential Task Force on Regulatory Relief chaired by Vice President George Bush.

FDA Commissioner Frank Young was called in to make a presentation on proposals to speed up the drug approval process at a meeting of the Presidential Task Force late last month. At the meeting, Young laid out a plan which would eliminate extensive Phase III efficacy studies involving hundreds of people for drugs to treat life-threatening diseases.

Phase III studies are designed not only to prove a drug's effectiveness in a large patient sample, but also to monitor the drug for long-term side effects. Many have argued that it is more important to get life-saving drugs to patients quickly and take the chance that the drug may cause side effects than to wait and collect all the data while patients are dying.

Young's proposals would establish a mandatory conference between the FDA and company sponsoring the drug at the end of Phase I of the clinical trials, in which the drug's safety to humans and dose are evaluated in less than 50 patients.

At the post-Phase I conference, FDA would sit down with representatives from the company to go over proposed protocols for the Phase II studies — where the drug's effectiveness is tested for the first time in about 150 people — to make sure that the company understands what type of data the FDA is looking for. Phase II tests would be expanded to include enough patients to demonstrate efficacy without the need to move on to the more time-consuming Phase III tests. Once the drug is on the market, companies would be required to file 'Phase IV' data on the drug's side-effects, to enable the FDA to monitor the drug and to revoke its approval if side-effects proved to be severe.

The regulations differ from the 'treatment IND' exception created by FDA last year (see *Nature* 326, 536; 1987). Under the 'treatment IND', a company may sell drugs for serious or life-threatening diseases only if it demonstrates that it needs the money to continue the clinical trials

necessary to file an Investigational New Drug application for approval. The treatment IND exception was formulated to aid small drug companies — primarily biotechnology companies — who may be sponsoring a drug for the first time and have limited cash reserves. But so far, only five companies have taken advantage of the clause because of fears that it would prolong the review process by taking away the pressure on FDA to approve a new drug for use on patients.

The approval last year of AZT as a treatment for AIDS provided the impetus for the new clinical testing regulations. The Phase III testing requirement was dropped for AZT when the Phase II studies showed that it was indeed therapeutic. Because no alternative treatment was available and patients were not expected to live long enough to suffer long-term side-effects, the FDA allowed Burroughs Wellcome to sell the drug without the larger Phase III trials.

In response to pressures to act on drugs being developed for AIDS in a more timely manner, the FDA also established last year a fast-track '1AA' review process for AIDS therapies. But the new '1AA' designation simply ensures that the paperwork is handled faster once the clinical trial data are collected; it does not allow the sponsors of a drug to side-step any of the studies required to demonstrate a drug's safety or efficacy.

As Commissioner of the FDA, Young has been praised for increasing the communication between the agency and companies with drugs in the approval pipeline, but has been criticized for taking a hard line in requiring that a drug be tested adequately, especially in the cases of ribavirin for AIDS and tissue plasminogen activator (TPA) for heart attacks. Sidney Wolfe, of the Health Research Group of Ralph Nader's Public Citizen organization, calls the latest changes a "thinly disguised public relations ploy" to boost Young's image and to make Bush seem more sensitive to health issues as part of his Presidential campaign. Wolfe says the FDA is already empowered to carry out expedited drug reviews, as illustrated by the approval of AZT, without additional rulemaking.

Once the new regulations are drafted, a process expected to take several months, they will be published in the *Federal Register* and opened to public comment.

Carol Ezzell

Seal epidemic still spreading

London

THE epidemic that has killed 7,000 seals in the North Sea and the Baltic in the past six months remains a mystery to researchers from six European countries who took part in a two-day emergency conference organized by Greenpeace International in London last week.

No causal link has been established between the disease that has wiped out half the population of common seals in mainland Europe and two viruses that have been identified in the dead animals. And though it is thought likely that pollutants play a role in the epidemic, the researchers cannot yet say which chemicals are implicated or exactly what their role is. The working party of European researchers is recommending an international, multidisciplinary research project. Meanwhile, leakage into the sea of polychlorinated biphenyls (PCBs) and other persistent pollutants should be stopped until they are proved to be harmless. But questions arise as to the role of pollutants because the area in Denmark



A baby seal, surrounded by victims, Isle of Sylt, where the disease seems to have originated is relatively free from pollution.

The party's recommendations will be presented this week to a government-level meeting in Stockholm.

Symptoms of the disease vary, but those most prominent are lethargy, respiratory problems, skin lesions, abortion, diarrhoea and nervous complaints. Most of the affected seals die within two days and are found to have inflammation of the lungs, brain, liver, intestines and skin. Two viruses have been identified in the dead seals. One is a herpes virus, already linked to pneumonia outbreaks in seals. The second virus is of the picorna family. How the virus spreads and why the seals react so badly is still unknown, says Dr Minnie Courtney of Queen Mary College, London. There are clear indications that the seals' immune system is being suppressed, and although the link between pollutants and this epidemic is not yet proved, persistent organic compounds such as PCBs and dioxin are known to affect immune response. Christine McGourty

Is Stanford's false economy on collider costing dear?

Berkeley

AFTER more than a month of operation, the new linear collider at the Stanford Linear Accelerator Center (SLAC) is still not producing Z^0 particles, and SLAC has assembled a task force to tackle the problem.

Early trials demonstrated the feasibility of the new collider, which accelerates bunches of positrons and electrons in a linear accelerator, then diverts them through opposing arcs which focus them on a collision course. But problems associated with the collider's cost-saving construction have hampered full-scale operation.

When focusing the beam earlier this year, SLAC ran into trouble because of the irregular shape of the arcs, the result of construction that saved money by following the contours of the hillside (see *Nature* 333, 8; 1988).

The current problem is the inability to maintain stable, reliable delivery of elect-

ron and positron bunches to the arcs. SLAC spokesman Michael Riordan says the problems are due to a high failure rate of components in the 20-year-old linear accelerator.

Only 3 per cent of the operating time in July was suitable for data collection. The task force will attempt to identify changes necessary to raise that to 50 per cent or better.

Marcia Barinaga

■ Some experts have already expressed doubts about the wisdom of "cutting corners" at SLAC in order to compete with European rival CERN. Writing in *Nature* last September, a correspondent concluded that "even if [SLAC] is the first Z^0 factory to operate it could be that if LEP runs to specification, its superior luminosity will allow physicists at CERN to mop up all the exciting results" (*Nature* 329, 107; 1987). CERN's Large Electron-Positron Collider (LEP) is now in the final stages of construction, with the first collisions due next year. □

Another chance to reorganize India's science planning

New Delhi

MR Rajiv Gandhi, the Indian prime minister, has been asked to establish a National Science and Technology Commission (NSTC) to integrate science and technology into economic planning. This is one of the chief recommendations in the *Perspective plan for 2001 AD* drawn up by Gandhi's 11-member Science Advisory Council (SAC), created two years ago specifically to evolve a strategy for using science as a tool for economic development.

In the past, atomic energy, space, electronics and other agencies have pursued projects in isolation, and knowledge generated by them "is not adequately being utilized in the growth of the national economy", says the SAC report released last week. The document, to be discussed in larger scientific forums before its formal adoption by the government, is said to provide a new prescription for integrating science with economic planning.

NSTC is expected to identify technology missions and to integrate science into planning by coordinating the work of government and of private and public sector agencies concerned with technology development. The planning commission and NSTC will forge a relationship under which the former will set goals and the latter will supply scientific input to achieve them.

SAC's proposal for a separate commis-

sion is seen as a reflection on the science cell within the Planning Commission. At present headed by Professor M.G.K. Menon, who is also the prime minister's science adviser, the cell is supposed to provide the appropriate science inputs for planning. But in SAC's view, the cell merely allocates funds to projects already identified by scientific agencies without taking account of overall planning and relevance to national needs. "We are convinced", says Professor C. N. R. Rao, chairman of SAC, "that there is need for considerable change in the manner in which we use science and technology in the total planning process and the way we execute the plans."

Rao said the perspective plan, if implemented faithfully, will help to double food production and gross national product (GNP), arrest population growth, bring health and literacy to a vast majority, reverse erosion of ecology and make a major dent in housing and unemployment, besides making India a leader in a few selected areas of science and technology by the turn of the century. To achieve all these objectives, SAC estimates that India must invest 2-3 per cent of GNP in science and technology, compared with 1 per cent at present. Specific proposals include a new system of links between universities and national laboratories along the lines of the Centre National de la Recherche Scientifique in France, and a centre for

Call for scientific bric-à-brac

London

A UNIQUE and diverse international collection of scientific apparatus is being assembled by the British Science and Technology Trust for auction next year. And with the money raised, the trust will set up an international science centre for the blind and handicapped.

Professor Paul Cook of Brunel University, president and founder of the trust, has contacted more than 500 academics for donations, and is overwhelmed by the response. Items collected so far include the compass used by Sir Alec Rose when he navigated the world; the chair of Nobel prize winner Sir Peter Medawar; a piece of the purest platinum; a linear accelerator; early X-ray crystallography pictures that led to discoveries in genetics at the Lab-



What am I bid? Paul Cook offers a slide-wire potentiometer used by the MRC in 1920 to calibrate a linear accelerator.

oratory for Molecular Biology in Cambridge; and equipment used in early experiments in radiation physics. The trust is also hoping for an old Superfortress plane.

A synthetic diamond given to Sir Alan Cottrell, and some pure platinum, have also been donated and are to be made into a ring by Aspreys, the royal jeweller. One company, which has asked to remain anonymous, has suggested that industry collaborate to buy the ring for the prime minister, Margaret Thatcher. And the company, which estimates the ring's value at £250,000, has put forward £10,000. The trust hopes to raise the public profile of science and to interest young people. Any money raised over the £250,000 needed to set up the centre will pay for scholarships and visits to science institutions.

Christine McGourty

advanced study in basic science similar to that at Princeton, New Jersey.

Whether or not Gandhi will act on SAC's recommendations remains to be seen. Fourteen years ago, the late prime minister Mrs Indira Gandhi asked scientists to prepare a science and technology plan, but the two volumes produced by a committee after five years of labour is now gathering dust.

K.S. Jayaraman

Japanese plans for international seismometer network shape up

Tokyo

JAPANESE seismologists are hopeful that plans to establish an international network of highly sensitive seismometers in the western Pacific and east and south-east Asia may soon begin to take shape. Although it will probably be the early 1990s before the long process needed to gain full government backing can be completed, there is confidence that the project will gain momentum much earlier.

The network, called POSEIDON (Pacific Orient Seismic Digital Observation Network), was described at an international symposium at Tokyo University earlier this month, and is intended to form part of an international global network being set up under the aegis of the Federation of Digital Broadband Seismic Networks (see *Nature* 328, 105; 1987).

The latest digital seismometers are capable of detecting ground motions over a tremendous range of frequencies and amplitudes, from tiny microearthquakes with periods of a fraction of a second up to Earth 'tides' caused by variations in the gravitational attraction of the Moon and Sun. Networks of these seismometers are being deployed around the world and will eventually enable geologists to 'X-ray' the Earth from different angles and build up a detailed three-dimensional picture of the planet's internal structure.

The POSEIDON network will consist of about 50 stations in a grid extending from the Aleutian islands to Papua New Guinea, including about 20 on the floor of the western Pacific Ocean.

Until now these highly sensitive seismometers have been used only on land. Three options for sea-floor operation are being considered: temporary 'pop-up' systems that will float to the surface after a period of about six months on the sea-floor; permanent seismometers linked to a surface buoy that will be equipped with an antenna for satellite transmission of data to a central processing centre; and a network directly linked to Japan by submarine optical fibre cable.

The cable option is the most appealing, but it is by far the most expensive — ¥35,600 million (\$260 million) for 20 stations as opposed to a few thousand million yen each for the other two systems, according to estimates by T. Kanazawa of Tokyo University.

Another approach being considered by Tokyo University's Earthquake Research Institute is to connect seismometers to an old trans-Pacific telephone cable, TPC-1, which will be abandoned next year and replaced with optical fibre cables. But there are objections from the US military, who currently use a section of the cable.

Regardless of which approach is adopted, Duncan Agnew of the University of California foresees problems in placing the seismometers on the sea floor. The seismometers are highly sensitive to tilt and much of the Pacific Ocean is floored with soft soupy mud. But Robert Geller, an associate professor at Tokyo University, is confident this problem can be overcome at sites with firm substrates. Another problem for these supersensitive instruments may be background noise, both on the sea floor and on land.

Japanese seismologists at the symposium went to great pains to emphasize on the one hand that POSEIDON has yet to receive government backing (apart from a few million yen for a feasibility study) and on the other to appeal to scientists in countries covered by POSEIDON to join the project.

There are ways in which seismometers can be deployed before POSEIDON is officially launched. Developing countries can apply to the Japan International Cooperation Agency for overseas aid to install seismometers and for training of their scientists. And Japanese seismologists can (and almost certainly will) apply to the Ministry of Education, Science and Culture by this November for grants of a few million yen to buy seismometers that can be installed overseas in fiscal year 1989. According to a member of one of the POSEIDON working groups, scientists from Indonesia, the Philippines and the Soviet Union have expressed interest in initiating the project by these sort of routes.

But a bureaucratic obstacle stands in the way: the Japanese government has no rules for the permanent installation of scientific equipment overseas (see *Nature* 311, 5; 1984). But there is a way around the problem. The equipment can be 'temporarily' installed and then 'abandoned' under a bureaucratic procedure for dealing with obsolete equipment.

Harumi Aoki of Nagoya University, who is chairman of the Global Seismology subcommittee organizing POSEIDON, says it will probably take until the early 1990s to win government funding for POSEIDON. But many organizations linked to various ministries are already backing the project. If the cooperation of other countries can be obtained the subcommittee will be in a much better position to apply to the government for support.

Robert Geller told foreign participants that they should not be discouraged if it takes five years to launch POSEIDON, because in Japan once a project gets going it's very hard to stop. Once the flow of funds starts it will continue for many years.

David Swinbanks

Space station surprise

PROSPECTS for the space station took an unexpected leap forward last week when the US Congress agreed to cut back on funds for the Environmental Protection Agency and disaster relief and increase support for the space station. The project will receive \$900 million in a new appropriations bill passed by both houses of Congress. The bill, which now goes to President Ronald Reagan for signing, provides any \$250 million less than NASA originally requested and \$700 million more than a Senate bill considered just last month.

A.A.

Venereal disease spinoff

THE West German Health Ministry last week reported a drastic reduction in the number of cases of sexually transmitted diseases. The incidence of gonorrhoea and syphilis in the first half of 1988 was one-quarter of that reported in the first half of 1985. The ministry attributed the reduction to the AIDS information campaign. The number of newly registered infections with human immunodeficiency virus (HIV), the virus that causes AIDS, has also stopped increasing in West Germany. Until now, 20,956 HIV-positive people have been registered, among them 2,939 women.

S.D.

Shuttle success at last

AFTER delays due to a leaky fuel line and a sticky valve, the main engines of the US space shuttle Discovery were successfully fired for 22 seconds on 10 August. The test caused great rejoicing within NASA, which has had little to cheer about lately, but before Discovery can be launched there remains one final test of the main booster rockets, presently scheduled for today, 18 August, in Utah, and the leaking fuel line must still be fixed. NASA officials now believe they can make the repair without rolling Discovery back off the launching pad, and are talking of a launch in late September, although no date has officially been set.

D.L.

Forward march

THE space programme of the People's Republic of China received a boost on 5 August with the launch of a Long March 2 rocket carrying West German microgravity experiments. The payload was to remain in space for eight days and return to Earth by parachute in China. A set of 104 protein samples from West German pharmaceutical companies was to crystallize in space under near-weightless conditions. The companies will use the crystals to study the structure and function of the compounds once they return to Earth.

The DM1.2 million launch costs and 80 per cent of the DM1.5 million project costs were paid by the West German Research and Technology Ministry. The pharmaceutical companies paid the rest.

S.D.

French research and development gets the boost it was promised

Paris

As is the custom, French ministers have been told before the summer recess how much money they will have to spend next year. While some ministers will be disappointed, minister for research and technology Hubert Curien started his vacation a happy man, with FF3,000 million (\$484 million) more in his purse for 1989 than for the current year. This increase, of over 7.6 per cent, is way ahead of inflation (2.6 per cent) and confirms the new government's determination to give high priority to research and higher education.

An increase in spending was certain. Shortly after the May elections, Curien was given part of a FF1,200 million 'advance' to pull the academic community out of the doldrums it felt itself to be in as a result of the previous government's cuts (see *Nature* 333, 584; 1988). Under Prime Minister Jacques Chirac, research spending in 1987 and 1988 was allowed to fall in real terms, and jobs were lost, with a promise that the private sector would be helped to make up the difference.

The new Prime Minister Michel Rocard has been given full rein not only to reverse this downward trend, but to try to approach the 3 per cent of gross domestic product (GDP) for research that François Mitterrand once promised by the end of the decade. With current spending at 2.26 per cent of GDP, France lags slightly behind the United Kingdom and West Germany and it now seems unlikely that the target will be reached by 1990.

Curien, who was given his own ministry in Rocard's second government, has also regained control of the research budgets of some other ministries, such as industry. This combined civil research and development budget (BCRD) is felt to allow better management of the nation's civil research spending, but was abandoned by the previous government, which lumped defence and civil research spending together. It is the BCRD that has increased by 7.6 per cent, to FF42,300 million.

Although the 1989 budget reflects Mitterrand's election promises, with education also receiving a FF11,000 million (5.5 per cent) boost, it is considered a 'one-off.' The government wants, next year, to reinstate the socialists' previous policy of long-term, statutory budget programmes for education and research which will underwrite a net annual progression. This, they hope, will avoid the 'accordion' effects of boosts and cuts experienced with changes of government over the past ten years. This has particularly affected higher education and research jobs where substantial new recruitments are now necessary to offset a

spate of retirements in the early 1990s.

If job creation is one of Curien's priorities, with 900 new researcher, technician and administrator posts expected for 1989 — compared to cuts last year — stronger links with industry are another. Earlier this month, Curien unveiled a programme, with FF120 million to spend this year, for collaboration between industry and public sector research. The programme aims to stimulate 'technology leaps' in French industry, with purification and manufacture of recombinant proteins one of the two priorities named so far. **Peter Coles**

South Africa keeps them guessing

Oxford

A FURTHER round of talks on the possibility of South Africa signing the Nuclear Non-Proliferation Treaty ended inconclusively in Vienna last week. South Africa's Minister of Foreign Affairs Pik Botha met Soviet, US and British officials for discussions before the general assembly of the International Atomic Energy Agency (IAEA) next month, where there is almost certain to be a renewed bid to expel South Africa (see *Nature* 334, 464; 1988). Botha stated afterwards that the South African government was still considering the implications of signing the treaty.

Botha also took the opportunity to engage in some characteristic sabre-rattling, declaring that his government had "the capability to make the bomb should we want to". When asked if South Africa already had nuclear weapons, he declined to enlarge on his statement. That South Africa has weapons capability is no major revelation, although this is the first time that it has been suggested at cabinet level by a member of the present government. Botha's statement differs from one made by former South African prime minister John Vorster in 1976 only in that it omits a rider about the country's intentions. Vorster said: "We are only interested in the peaceful applications of nuclear power. But we can enrich uranium, and we have the capability. And we did not sign the nuclear non-proliferation treaty."

Botha claimed further that the IAEA director-general, Hans Blix, is biased against South Africa over curbs on the spread of nuclear weapons. "I believe that Dr Blix doesn't want us to be part of the agency", he declared. Blix has denied this, saying that the agency would inspect South African nuclear facilities in exactly the same way as it checks plants in other member states **Michael Cherry**

Towards caring not curing

London

A REVOLUTION in medical education was set in motion last week at a five-day conference in Edinburgh, Scotland, organized by the World Federation for Medical Education (WFME). The outcome of the conference, a strategy document called The Edinburgh Declaration, outlines radical changes in training in medicine. The declaration recommends more emphasis on community health and preventative medicine in a reversal of the present bias towards hospital-based training and curative medicine. In a pre-conference document, the WFME says that as teachers are becoming more specialized, medical curricula are becoming increasingly compartmentalized; and training, instead of being focused on the community is focused on the hospital bed. As a result, "It is not surprising if medical students emerge with great regard for high technology and the most recent knowledge, low regard for less dramatic elements of health promotion, prevention of disease and caring rather than curing".

The conference was held in collaboration with the World Health Organisation and the United Nations Children's Fund among others, and was attended by representatives of universities, ministries of health and education and funding institutions such as the World Bank and the Rockefeller Foundation.

The declaration emphasizes that since there will have to be legislative changes, political will is necessary for reform to succeed. "The endorsement of the declaration by governments is 'absolutely vital'", says Walton. The first of six consultations with ministers will be a European conference in Portugal in November. These consultations should all be completed before the declaration goes before the World Health Assembly in 1989.

The cost of the reforms will be met by international funding agencies and coordinated via a new institute in Edinburgh. But Walton says no difficulties of cost are envisaged; nor of infrastructure. The only obstacle to implementation of the reforms, he says, will be the "entrenched attitudes of medical institutions".

But change will come about since it is being instigated from within the profession and not forced on it from without. "Everybody has wanted these reforms for decades," says Walton. And he says half of them could be introduced immediately.

The new director-general of the World Health Organisation, Dr Hiroshi Nakajima, pledged his full support for the declaration. "Technological excellence must not be an end in itself for the medical profession," he said. **Christine McGourty**

Tests for new AIDS treatment begin in three clinics

Washington

CLINICAL trials of recombinant soluble CD4 receptor produced by the biotechnology company Genentech as a treatment for AIDS began last week at three centres in the United States. Human soluble CD4 is the first genetically engineered drug to be tested against AIDS.

CD4 is the receptor carried on the cell membrane of T-lymphocyte cells in the blood to which the human immunodeficiency virus (HIV) binds. Genentech produces soluble CD4 by expressing the DNA which codes for the extracellular portion of the CD4 receptor in mammalian cells. The company has shown in a study published in *Science* (238, 1704; 1987) that the resulting truncated form of CD4 binds to the envelope protein of HIV *in vitro*, and neutralizes virus infectivity.

In vitro studies have also been published using soluble CD4 produced by Biogen, Dana-Farber Cancer Institute, SmithKline and French, and the Basel Institute of Immunology in Switzerland (*Nature* 331, 76–86; 1988), and by Ortho Pharmaceuticals (*Science* 240, 1335; 1988). The California biotechnology company Gene-Labs is also producing recombinant CD4. But all of these groups are lagging

behind Genentech, and have not yet filed for approval from the US Food and Drug Administration to begin human clinical trials. Biogen hopes to seek approval to enter the clinic this autumn.

Two separate clinical trials of Genentech's soluble CD4 have now begun: one headed by Samuel Broder at the US National Cancer Institute, and the other by Jerome Groopman at the New England Deaconess Hospital in Boston, Massachusetts, and by Paul Volberding at San Francisco General Hospital, in conjunction with the US National Institute of Allergy and Infectious Disease's AIDS Program. Approximately 50 AIDS and AIDS-related complex patients will be enrolled in the six-month studies, which will determine whether the treatment is safe and establish the dose range.

The side effects and efficacy of soluble CD4 in the body are unknown. Theoretically, patients could produce antibodies to the soluble CD4 that could then turn on the CD4 molecules on their own T-cells, producing severe immune problems. But immune system side effects predicted for recombinant vaccines for AIDS have not developed, and none are expected for soluble CD4.

Carol Ezzell

Doubts over long-term space flight after cosmonaut's death

London

THE death last week of cosmonaut Anatolii Levchenko, less than eight months after his 8-day mission aboard the Mir orbital station, has inevitably aroused speculation concerning the Soviet shuttle programme. The launch of the Soviet reusable spacecraft, is expected within the next few months and the launch date should be unaffected. Levchenko was one of the two test-pilot/cosmonauts training to test-fly the craft — the other is Igor Volk, but the first flight of the shuttle is in any case to be unmanned.

The doctors have been adamant that Levchenko's death was not connected with his space mission, from which, post-mission reports indicated, he had adapted with remarkable ease. The cause of death, given in the official obituary as "a grave illness", has since been amplified to a brain tumour. If, however, this tumour was not identified in the rigorous medical checks which Levchenko must regularly have undergone during his almost 20 years' career as test-pilot and trainee cosmonaut, the question must be asked: do any other cosmonauts in the Soviet space programme have health problems which

have gone undetected? This could have considerable implications, not only for the Soviet Mars programme, but also for Mir and future orbital stations.

During the past few years, successive record-breaking sojourns in zero gravity have been clearly related to an eventual Mars mission. A permanently occupied orbital station has, from the beginning, been a staple of all Soviet space plans, but this could be staffed on, say, a six-week shift system. Long-stay cosmonauts, however, test out the medical and exercise programmes which will be essential to a Mars flight. But any such mission presupposes that the cosmonauts undertaking it will be initially in perfect health. Had Levchenko's destination last December been not Mir, but Mars, he might have died in mid-mission.

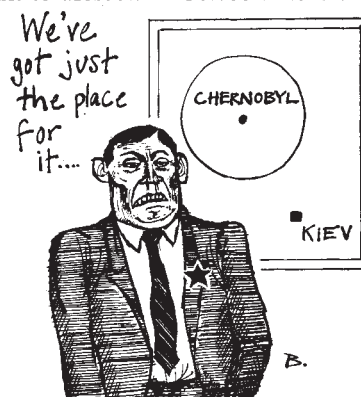
Soviet public opinion is already concerned at the cost of a Mars mission. *Sovetskaya Kultura* recently published a letter from a Moscow engineer, suggesting "those who are waiting in the queue for housing because they live in under 5 square metres per person, and those who spend their days in rooms with five or more others in old people's homes . . .

Nuclear waste may go to East

Munich

THE West German government is considering an offer by the Soviet Union to reprocess and permanently store West German nuclear waste. The offer came in the form of a letter to Chancellor Helmut Kohl delivered earlier in the summer by a West German official who had been visiting Moscow.

Government spokesman Herbert Schmülling does not expect the subject to be discussed during Kohl's forthcoming visit to Moscow in October. Nevertheless,



he said that the letter is being "processed" in the office of the chancellor.

The West German Environment Ministry, which is responsible for much of the nuclear programme, says that it stands by its "integrated waste-processing concept" which would ultimately have all nuclear waste from West German nuclear power plants processed and stored in West Germany. Schmülling said that West Germany still plans to complete its own reprocessing plant at Wackersdorf and permanent waste storage site at Gorleben despite "considerable opposition". The Environment Ministry confirmed that cooperation with the Soviets is possible. West German nuclear fuel rods are at present reprocessed at a French facility at La Hague. West Germany has other agreements with Great Britain and Sweden, both of which have announced cuts in their nuclear programmes in recent weeks.

Steven Dickman

people from Chernobyl should be asked their opinion of the cost. And if the Mars mission is abandoned, and Mir, perhaps, switched to a short-shift system, then the next step could be pressure to abandon the concept of continuous occupancy of the station. Much of the astrophysical and Earth observation work now carried out by cosmonauts could well be done by remote sensing satellites. For those in the Soviet space programme who fear the new pressures of public opinion, Levchenko's death must seem more than a purely human tragedy.

Vera Rich

DoE's management of US nuclear reactors still not satisfactory

Washington

DESPITE big strides in improved safety performance, a National Research Council report* concludes that there is still scope for improvement in the management by the Department of Energy (DoE) of its nuclear reactors. The report, released last week, appears just one week after DoE responded to a recommendation of an earlier Research Council report† by proposing two new production reactors to restore US capacity to produce tritium for nuclear weapons.

The two research council reports were originally requested by DoE following the Chernobyl nuclear reactor accident in January 1986. The first report focused on the defence production reactors — three heavy water reactors at the Savannah River Plant in South Carolina, and the graphite-moderated N Reactor at the Hanford Nuclear Reservation in Washington State. In addition to specific technical safety concerns raised in large measure by the age of the reactors, the 1987 report criticized DoE for a lack of attention to safety, and urged the agency to adopt a clearly articulated safety policy.

DoE has since taken several steps to improve its safety performance. The N Reactor has been shut down, as has one of the three Savannah River reactors, and the remaining two are operating at decreased power. There has been a 50 per cent increase in the staff of the Office of Environment, Safety and Health, as well as improved administrative procedures and more frequent and broader safety appraisals. The agency also created an independent Advisory Committee for Nuclear Facility Safety, chaired by John F. Ahearne, on which it will "rely heavily".

But Richard Meserve, chairman of the panels that produced the two reports, says the Ahearne committee alone is not sufficient to resolve safety concerns. He says a committee that meets one or two days each month can only begin to address the type of safety issues that operating high powered nuclear reactors present.

The most recent report examines the five Class A (20 MW or greater) DoE test and research reactors. Again ageing is a problem, as is the diversity of designs used in the reactors: the Advanced Test Reactor is a 250 MW light water reactor; the Experimental Breeder Reactor II and the Fast Flux Test Facility are both liquid metal cooled reactors; the High Flux Beam Reactor is a heavy water reactor, and the High Flux Isotope Reactor uses

light water as a coolant and beryllium as a reflector. The multitude of designs makes numerous safety analyses necessary.

Acknowledging the need to replace the Savannah River reactors, DoE announced last month its plan to build two new production reactors; a heavy water reactor at Savannah River and a modular high-temperature-gas reactor (MHTGR) at the Idaho National Engineering Laboratory. The two-reactor approach to solving the tritium supply problem eases security and political concerns, but the \$6,800 million cost of pursuing both technologies may not be fiscally acceptable to Congress.

In January this year, Energy Secretary John S. Herrington requested the DoE Energy Research Advisory Board (ERAB) to produce an analysis of four options being considered by DoE to replace the Savannah River reactors: a new heavy water reactor, a light water reactor, a liquid metal reactor and the MHTGR.

In a report released last month‡, ERAB concluded that although all the proposed designs could supply US tritium needs, the heavy water reactor is the most mature technology and would present the fewest technical hurdles. The existence of personnel and facilities at Savannah River,

would help keep capital costs down.

But ERAB further concluded that a multiple reactor strategy should assure a steady production capability, reduce uncertainties in the new production reactor schedule and "minimize the technical risks to national security". After heavy water reactors, ERAB considered tritium production technology to be most advanced with MHTGR designs.

Although a heavy water reactor is the most cost-effective way of producing tritium in the short run, proponents of alternative technologies have pointed out that other designs, notably liquid metal and MHTGR, have commercial applications that could help to offset development costs. But ERAB warns that it has been traditional US policy to separate military and civilian nuclear reactor use. "A change in this policy could introduce political risks for the new production reactor project." It could also harm prospects for civilian nuclear power, according to the ERAB report.

DoE must now produce environmental impact statements on its proposals. It will also seek design advice from its newly created Advisory Committee on Nuclear Safety. DoE has \$60 million in its 1989 budget for work on the new production reactor, but the tough funding decisions — especially as they relate to proceeding with two reactors — will probably wait until after the fall elections.

Joseph Palca

Universities a forgotten resource

THE United States is in danger of losing a national resource if it fails to take steps to restore the vitality of university research reactors (URRs), according to a report§ released last week by the National Research Council.

There are at present some 40 reactors operating on campuses around the country. The most powerful is a 10 MW reactor at the University of Missouri Columbia campus, but there are two dozen reactors that produce less than 250 kW of power, right down to a reactor at Manhattan College rated at 0.1 W.

Many of the reactors were built 20–30 years ago, when the now-defunct Atomic Energy Commission was spending heavily on developing nuclear power. But when the commission's responsibilities were divided between the Nuclear Regulatory Commission and the Department of Energy (DoE), the URRs fell through the cracks. Today, DoE contributes about \$2 million for the operation of the reactors, mainly in the form of fuel.

Operating costs, even of fairly low powered reactors, are typically hundreds

of thousands of dollars a year, and a large facility such as Missouri has an annual budget of nearly \$4 million.

Funding comes primarily from indirect charges against research grants, direct support from universities and 'service' functions. These functions can in some cases provide a substantial return. Missouri has been using its neutron flux to irradiate topaz, causing the gemstone to take on a blue colour that increases its commercial value.

But without coordinated federal support, URRs face a difficult future. David Shirley, director of the Lawrence Berkeley Laboratory and chairman of the panel that produced the report, says he believes that an interagency advisory panel would help to coordinate federal support and that a URR would form an excellent focus for one of the new science and technology centres that have been proposed by the National Science Foundation.

The report urges DoE to adopt a network strategy like that centred at the Institut Laue-Langevin in Grenoble, France, where a 57 MW reactor acts as a centre for nuclear research activities, and maintains close ties with university reactor facilities.

Joseph Palca

* *Safety Issues at the DOE Test and Research Reactors*. National Academy Press, Washington, DC, 1988.

† *Safety Issues at the Defense Production Reactors*. National Academy Press, Washington, DC, 1987.

‡ *Assessment of Candidate Reactor Technologies for the new Production Reactor*. A report of the Energy Research Advisory Board to the United States Department of Energy. Washington, DC, 1988.

§ *University Research Reactors in the United States — Their Role and Value*. National Academy Press, Washington, DC, 1988.

CORRESPONDENCE

Evidence of non-reproducibility

SIR—The paper by Benveniste and colleagues¹ suggesting that the presence of a molecule in aqueous solution can transfer its properties to the solution in the absence of the molecule itself has, as the authors surely anticipated, caused considerable scepticism in our laboratory and others. Most new and revolutionary ideas are initially met with scepticism. But the essence of science is the universality of the principles, and if the ideas presented in this paper are true, they should extend to similar systems.

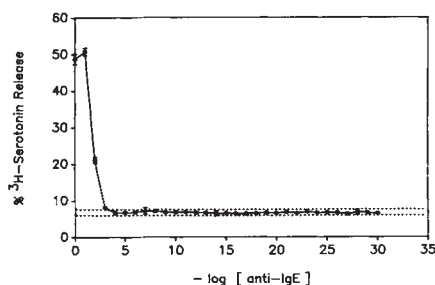
We have therefore attempted to reproduce the results in a system that models the allergic response. Cells of the cultured rat mast cell tumour line RBL-2H3 were primed with IgE and loaded with 5-hydroxytryptamine (serotonin), and assayed as previously described², with the modifications described in the figure legend. The figure (representative of three similar experiments) shows that above a dilution of 1×10^4 of the antibody, there was no secretion above the spontaneous release (cells incubated in Hanks' buffer).

We cannot explain the discrepancy between these results and the results presented by the Benveniste group. The RBL-2H3 system we describe is a more

clearly defined, and probably more sensitive system, in that it uses a pure population of cells in large number (500,000 cells per assay as compared with the 60–120 cells counted per assay in the basophil system), a more quantitative assay method (release of labelled serotonin rather than loss of toluidine blue staining cells), a clearly defined IgE, and a purified antibody preparation instead of whole antiserum as reported by Benveniste and colleagues (although the claim is made that similar results were obtained with monoclonal anti-IgE, antigen, phospholipase A2, and Na^+ and Ca^{2+} ionophores). One other, presumably trivial, difference between our study and the previous one are our use of Hanks' buffer instead of Tyrode's. We therefore conclude that, despite the elaborate controls of the previous study, the results do not describe a new scientific principle, but instead must represent, at best, some peculiarity of the assay, and at worst, an intriguing artefact.

JEANCLARE SEAGRAVE

Cell Pathology Laboratory,
Surge Building,
University of New Mexico,
School of Medicine,
Albuquerque, New Mexico 87131, USA



RBL-2H3 cells were primed with $1 \mu\text{g ml}^{-1}$ monoclonal anti-dinitrophenol mouse IgE and loaded with [^3H]serotonin ($3 \mu\text{Ci ml}^{-1}$) for 3 h. The cells were washed and resuspended in modified Hanks' buffer (124.5 mM NaCl, 4.94 mM KCl, 0.4 mM KH_2PO_4 , 0.31 mM Na_2HPO_4 , 15.2 mM NaHCO_3 , 10 mM HEPES, 1.7 mM CaCl_2 , 0.73 mM MgSO_4 , 5.45 mM glucose, and $0.5 \mu\text{g ml}^{-1}$ bovine serum albumin, pH 7.4) at 5×10^6 cells per ml. Tenfold dilutions of an affinity purified polyclonal anti-mouse IgE from rabbit serum were prepared in the same buffer, with vortexing for 20 seconds at each dilution as described in ref. 1. One hundred microlitres of the diluted antibody and one hundred microlitres of the cell suspension were combined in an Eppendorf microfuge tube, and incubated for 20 min at 37°C . The secretion was stopped with 0.5 ml of ice-cold phosphate-buffered saline and the cells were pelleted by centrifugation for 2 min in an Eppendorf microfuge. The serotonin in 0.5 ml of the supernatant was determined by scintillation counting. Assays were performed in triplicate. Error bars are shown where the error exceeded the size of the symbol. The dotted lines indicate the range of spontaneous release (buffer alone), measured at several times during the assay.

SIR—We have repeated the experiment reported by Davenas *et al.*¹, except that we used histamine release from human basophils rather than basophil degranulation to evaluate the biological effect of a highly diluted anti-IgE, because in several laboratories histamine release is the method of choice. (Benveniste and co-workers themselves, in describing their test, state that basophil degranulation gives results that correlate well with those of histamine release³.)

Circulating basophils from a good histamine releaser were separated on a dextran gradient and incubated with increasing dilutions (from 10^{-1} to 10^{-45}) of a goat anti-human IgE (anti-Fc, Sigma Chemical, St Louis). Experimental conditions were as described by Davenas *et al.*, including their protocols for shaking the tubes and changing pipette tips. Histamine released in the supernatants was assayed by the automated fluorimetric method³, and expressed as per cent of histamine release of the same aliquots of basophils treated with 0.4 M perchloric acid. As expected, a good histamine release was observed between 10^{-1} and 10^{-4} dilutions of anti-IgE, with a dose-response curve ranging from 53 per cent to 6 per cent of histamine release.

However, no histamine release was observed for any further dilution. Repeat-

ing the experiment gave highly reproducible results.

We therefore conclude that the findings of Davenas *et al.* cannot be extended to another technique. Should we speculate why water can remember something on some occasions and forget it on others? As your editorial states: "It will be time for celebrations of that kind only when a lot more water has run underneath this bridge". . . and on laboratory desks.

SERGIO BONINI
EMILIO ADRIANI
FRANCESCO BALSANO

Andrea Cesalpino Foundation and
Institute of Clinica Medica 1st,
University of Rome "La Sapienza",
00161 Rome, Italy

1. Davenas E. *et al.* *Nature* **333**, 816 (1988).
2. Benveniste, J. *Ouest Med.* **30**, 467 (1977).
3. Siraganian, R.P. *Analyt. Biochem.* **57**, 383 (1974).
4. *Nature* **333**, 787 (1988).

SIR—So now at last confirmation of what I have always suspected. Papers for publication in *Nature* are refereed by the Editor, a magician and his rabbit.

KEITH SNELL

Department of Biochemistry,
University of Surrey,
Guildford, Surrey GU2 5XH, UK

Regulating IVF

SIR—In the report "International outlook for embryo research" (*Nature* **333**, 791; 1988) there were a number of misrepresentations regarding the complex regulatory climate in Australia.

There is no *in vitro* fertilization (IVF) legislation as such in Australia, since medical issues are decided state-by-state, and not on a Commonwealth basis. Only two states have passed laws regarding IVF, namely Victoria and South Australia. In the other states, IVF is governed by self-regulation under the auspices of the National Health and Medical Research Council. The only stipulations are that embryos should not be kept in culture beyond 14 days, and that local institutional bioethics committees should authorize the work.

The Victorian IVF law had been passed by the parliament in that state in 1984, but the vital section of the law dealing with embryo research was not proclaimed until 1 July 1988. According to the now valid law, research can only be performed with spare embryos left over in IVF treatments of infertile patients. Moreover, such research can only be conducted if prior approval had been obtained from a review committee set up by the Health Minister of Victoria.

Between its formulation in 1984 and proclamation in 1988, the original law was amended in an important way. As the law now stands, up to the stage of syngamy

when the male and female nuclei fuse, the zygote is not regarded as an embryo. This means that the prohibition of creating embryos for research does not apply to the first 20 hours. Dr Trounson and his colleagues at Monash University in Victoria have requested this amendment from the minister on the basis that the 20-hour gap will give them sufficient time to test whether micro-injection of deficient sperm into the egg results in a chromosomally defective zygote.

Based on these tests, Trounson and his colleagues plan to go ahead and insert micro-injected eggs into the uterus of infertile women (or of fertile women in the case of infertile partners). Indeed, they have already done this earlier this year. As the law stands, the insertion of micro-injected eggs into the uterus of a patient in an IVF programme is not regarded as research, and hence does not need the approval of the Victorian IVF review committee. It is assumed that premature work would be stopped by the local bioethics committee. In the micro-injection case just mentioned this had not happened and so the work had gone ahead.

Clearly embryo research laws are very difficult to formulate in such a way that they encapsulate all that a parliament may wish to have regulated. It is hoped that the less than satisfactory legal situation in Victoria will serve Britain as a useful lesson for how not to draft such legislation.

DITTA BARTELS

*School of Science & Technology Studies,
The University of New South Wales,
PO Box 1, Kensington,
New South Wales, Australia 2033*

German angle

SIR—The leading article regarding the Embryo Protection Law being formulated by the West German Justice Ministry (*Nature* 333, 787; 1988) was loaded with subjective comments, which are indeed matters of opinion and not incontrovertible. To take them in order:

- It is indeed a "paradox that research should be singled out" for this legislation. Clearly this does not prove that the legislation is unnecessary. It may just need to be extended in scope.

- Our understanding of the ethical implications of medical research lags far behind our technological progress. A degree in biochemistry is not a qualification in moral philosophy, and the idea that researchers should or can "police themselves" is ill-founded at best, especially as profit, prestige and prizes are available to motivate those who break the rules.

- I disagree that it is small consolation to the research community for "history and the legal system" to provide restraint to scientific research. Civilized society must look to the mistakes of the past and protect itself from making the same mistakes

in the future. Research in a society devoid of such constraints would be intolerable.

- The article concludes by implying that the West German people should "appreciate that embryo research may, in due course . . . rid people of undignifying genetic diseases". We are a long way from curing anyone of a genetic disease, but have already begun ridding ourselves of PEOPLE, with what we consider "undignifying genetic diseases", through abortion and euthanasia. How similar that sounds to the rationale used by the Nazis to attempt the genocide of the Jewish race.

It is my sincere hope that generations to come will not have cause to condemn us for crimes which they view in the same way as we now view the crimes of Nazi Germany.

KENNETH W. M. COCHRAN

*AT&T Bell Laboratories,
200 Laurel Avenue,
Middletown,
New Jersey 07748, USA*

Moratorium call

SIR—We read with deep concern the letter by Braude *et al.* (*Nature* 332, 459–460; 1988) regarding the activation of the human embryonic genome. We feel that in this study the ethical bounds of scientific work have been overstepped.

Although biologists have developed techniques to manipulate mammalian embryos, the abuse of these techniques through experiments with human embryos (or pre-embryos, if one considers a preimplantation embryo not to be an embryo) must be condemned by the scientific community. Scientists cannot wait for public or political pressure to define the limits of their work. The time has come for an Asilomar-type moratorium to establish ethical boundaries for work on human embryos.

In our opinion this study is unethical. We think therefore that experiments of this kind should not be encouraged by their publication in any journal, including *Nature*.

(This letter represents the views of the undersigned senior scientists or graduate students, and should not be construed as necessarily representative of the views of the institute or the society for which we work.)

RUDI BALLING, KAMAL CHOWDHURY,
URBAN DEUTSCH, SUSANNE DIETRICH,
UWE DRESCHER, ULF HENSELING,
BIRGIT JOSTES, GUNNAR-INGI
KRISTJANSSON, TINE DE MAEYER,
ANDREAS PÜSCHEL, HANS SCHÖLER,
DARIA SIEKHAUS, FRANZ THEURING,
CLAUDIA WALTHER, ANDREAS ZIMMER

*Department of Cell Biology,
Max-Planck-Institute of Biophysical
Chemistry,
Göttingen, FRG*

Conservation plan

SIR—I was interested in your leading article (*Nature* 333, 284; 1988) outlining the problem of OECD food subsidies.

A possible useful way of harnessing the excess production would be to give or sell the food to developing countries at a reduced rate in return for land/forest conservation programmes. A scheme involving third world debt/conservation exchange seems to be working well.

The idea has a lot going for it. For example, the OECD farmers would be working for the retention of the remaining virgin land on this planet instead of purely for government hand-outs.

R. G. H. COTTON

*25 Avenue Athol,
Canterbury,
Victoria 3126, Australia*

New OSHA standards

SIR—You report (*Nature* 333, 590; 1988) that OSHA has set new permissible exposure limits for more than 400 chemicals. In fact, the agency has *proposed* new limits.

We are currently holding hearings on the proposal which should be completed by mid-August. We hope to issue a final standard for these chemicals before the end of the year.

AKIO KONOSHIMA

*US Department of Labor,
Occupational Safety and Health
Administration,
Washington, DC 20210, USA*

Ozone layer

SIR—Why not pressurize spray-cans with ozone?

EUGENE LEEB

*14 Salisbury Court,
Salisbury Avenue,
London N3 3AH, UK*

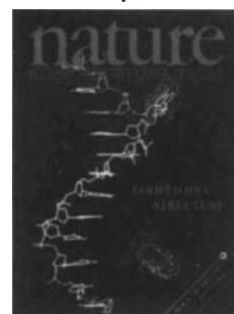
Cover blown

SIR—Wordsworth wrote: "Little we see in Nature that is ours". Indeed, the picture appearing on the cover of the issue of 7 July, attributed to us, is not our work.

ARTHUR M. LESK

VICTOR I. LESK

2 Field Way, Cambridge, UK



The figure used on the 7 July cover was in fact made by Leemor Joshua-Tor & Joel L. Sussman using the program FRODO, written by T. A. Jones and implemented on the PS890 by J. W. Pflugrath, M. A. Saper, J. S. Sack & F. A. Quiocho.

Statistical mechanics by numbers

Molecular dynamics is proving a useful technique well away from the applications for which it was originally devised. A guiding principle is to make a virtue of simplicity.

MOLECULAR dynamics is clearly well on the way to being a universal tool, as if it were the differential calculus. So much is plain from the range of problems now being tackled, and sometimes at least partially solved, by the application of this technique. Gone, it seems, are the days when people's ambitions in the field were restricted to the calculation of the properties of smallish molecules, with only distant hopes of being able to tackle the problems of, say, protein molecules. Now, molecular dynamics is being used to tackle problems which are quite general.

A recent issue of *Physical Review Letters* (25 July), for example, contains accounts of applications of the technique to the solution of the Navier-Stokes equation on the one hand and the surface melting of aluminium crystals on the other. The statement that all this has been made possible by the arrival of computer power in very substantial amounts is easy, almost banal, but it is only half the truth. The kinds of tasks now being taken up demand an intriguing degree of ingenuity, not least that of spotting what particular formulations of problems will yield results in keeping with physical intuition.

Yet in principle, molecular dynamics is as simple-minded approach to physical problems as could be conceived. Statistical-mechanics-by-numbers would be an appropriate name. For example, to calculate the thermodynamic properties of a molecule species (say methane) one has merely to construct a model of a single molecular, set it off in a particular state of motion and then use the appropriate laws of motion (classical or otherwise) to calculate succeeding configurations.

A few elementary simplifications are evidently possible: the translation of the centre of gravity of a molecule, or its rotation about that centre of gravity, can be separated (and temporarily put aside) from the outset. The task is merely to calculate succeeding configurations of the molecule, when the ergodic theorem (suggesting that a system in motion will in the long run traverse the whole of accessible phase space, but with a frequency proportional to the occupancy of different elements of that phase space in real life) can be used to yield proxies not merely for macroscopic quantities such as energy and entropy, but estimates of the time spent by a molecule in some specified state.

If the principle is simple, the practical difficulties are formidable. Arranging that molecular systems containing some hun-

dreds of particles should be calculable is not as great an impediment as that of arranging that the method of calculation is not a recipe for accumulating and then multiplying error from one step to the next. But there is evidently no reason why the same techniques should not in principle at least be applied to systems other than simple molecular systems. Indeed, in some respects, three-dimensional systems such as liquids may be more tractable than molecules, if only because the need to specify the force-law governing the interaction between each pair of particles may be enormously simplified for a homogeneous system, for example, there is at most one force-law to worry about.

The problem of melting is a good place at which to start. P. Stolze and J. K. Norsø from the Technical University of Denmark at Lyngby and U. Landman from the Georgia Institute of Technology (*Phys. Rev. Lett.* **61** 440; 1988) begin by remarking on the conflict of evidence and inference on the question of whether the melting of a solid begins as a surface phenomenon below the bulk melting temperature which then spreads through the bulk of the solid as the temperature approaches the bulk melting temperature. One particular difficulty is that there is experimental evidence showing some such effect as much as 150 K below the bulk melting temperature, but that such an effect has been calculated only for simple pair-wise force-laws between pairs of interacting atoms which are plainly inappropriate for metals such as aluminium.

In this case, the trick is to find a realistic force-law to represent the interactions between pairs of aluminium atoms, which has been constructed from a model in which metal atoms interact with the remainder of the structure of which they are a part through the average electron density in the neighbourhood. The result is entirely pleasing, if expected: a parameter, called the structure-factor, for each layer near the surface, which is related to the coherence of Bragg scattering from such a layer, is found to decrease sharply as the bulk melting temperature is approached from below. Between 500 K and 800 K, the ten outermost layers of the model of solid aluminium lose their order (which is what the experiments suggest). Perhaps the most startling implication of this result is that it points to ways in which the bulk melting-point of real metallic solids might be calculated, hitherto beyond most people's ambitions.

The calculation of the dynamical properties of liquids (by M. E. Colvin, A. J. C. Ladd and B. J. Alder of the Lawrence Livermore National Laboratory, *Phys. Rev. Lett.* **61** 381; 1988) may be even more portentous. What this group had done is to simulate the flow of a fluid by "stripping standard hard-disk molecular dynamics to its barest essentials". In other words, the world in which elements of a fluid (not necessarily atoms) are dealt with as if they were billiard balls (or disks in two dimensions) colliding by Newton's laws is replaced by that in which directional homogeneity is done away with by means of the supposition that the elements of the fluid are rigid hexagons (in two dimensions), each with the same orientation.

One advantage of this representation of a fluid in motion is that it is relatively easy to enumerate the kinds of collisions allowed; in this model, ten classes of collisions suffice. Each collision turns a pair of particles, whose velocities are specified, into another pair of particles whose velocities follow from the enumerated collision rules. Another advantage is that the time evolution of such a model may be simulated according to the techniques of cellular automata: pairs of elements of the fluid within reach of each other interact so as to yield another pair of particles with different velocities. Calculating the movement of the fluid is then simply a matter of following the evolution of the pattern of elementary velocities. A further advantage is that the model is admirably tractable with the help of parallel computing.

But surely the most one might hope to learn from such an exercise is the degree of further development of computers that will be necessary before realistic problems can be handled.?

That is the surprise. Far from being an impossibly crude model, the stripped-down molecular dynamics model is a remarkably good representation of some aspects of the behaviour of a fluid. For example, the long-time behaviour of the velocity autocorrelation function for a fluid element (the correlation between present and some past velocity and therefore a function of the time interval between then and now) is well predicted even by this crude model. Since this function is central to, for example, calculation of fluid viscosity, the result may be important. But above all, practitioners in this field should not let the search for perfection destroy their capacity to solve problems.

John Maddox

Surface geometry

New polymers in old honeycombs

Alan L. Mackay

THE philosophers of Laputa in *Gulliver's Travels* were satirized by Swift as carrying sacks of geometrical models so that they could show clearly what they were talking about. This practice is more necessary than ever as the repertoire of geometrical objects used in the grammar of structure expands, although computer graphics does make the models more portable. The geometry of surfaces in three dimensions is still quite hard to apprehend. Our minds seem to be better at imagining networks and trees than perceiving the relationships of space dissected by surfaces. Elsewhere in this issue (*Nature* 334, 598–601; 1988), E.L. Thomas *et al.* describe the physical realization of complicated interlocking surfaces that have the property of minimizing their area relative to the volume they enclose.

A burst of new ideas in crystallography has been prompted by the appreciation that periodicity can appear at various levels in the hierarchy of real structures from single atoms to bulk materials. Where sufficient atoms are concerned, surfaces become significant, and a grammar of structure for organic and inorganic materials, parallel to that developing for proteins, is unfolding. (See Anderson, S. *et al. Chem. Rev.* 88, 221–242; 1988.)

Several new geometrical structures with plane or curved surfaces have recently been added to the five platonic solids and other polyhedra used to describe crystal, liquid-crystal and quasicrystal textures and structures. Some of these are periodic minimal surfaces such as those studied by Thomas *et al.* The surfaces may be dense arrays of atoms, equipotential surfaces of the electrostatic fields in crystals or non-periodic structures, or even the contours of the 'Fermi surfaces' used by

solid-state physicists to describe the momentum distribution of electrons in a metal.

In the best-known periodic minimal surface, carbon atoms joined tetrahedrally by covalent bonds give the open diamond network (Fig. 1). An identical parallel framework can be inserted in the interstices of this framework, so that the two networks are disjunct but completely

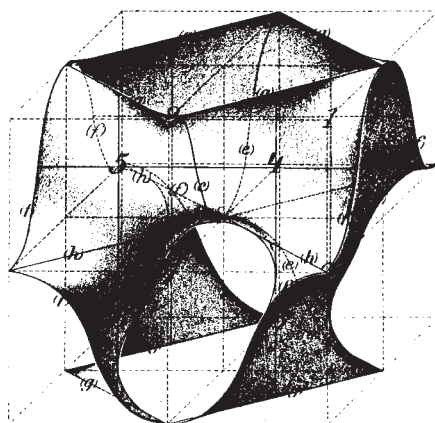


Fig. 2 Schwarz's own diagram of the unit tetragonal cell of the T-surface, to which the interface, in Thomas and co-workers' Fig. 2 (on page 600 of this issue) is related. (From Schwarz, H.A. *Gesammelte Mathematische Abhandlungen*, Springer, Berlin, 1890.)

interlinked. This arrangement occurs in cuprite (Cu_2O) and in high-pressure forms of ice (ice VII and VIII).

If we draw the interface half way between these two networks, we obtain a surface which divides the world into two congruent regions. This is known as the D (for diamond) or F (for face-centred cubic) surface of H.A. Schwarz (*Gesammelte Mathematische Abhandlungen*, Springer, Berlin, 1890). The precise mathematical definition of the surface requires that it should have everywhere 'constant mean curvature' — it is then called a minimal surface. Because the whole structure is periodic in space, it is called a periodic minimal surface.

A minimal surface is exemplified by a soap film with equal pressures on each side, where the two principal curvatures K_1 and K_2 along two perpendicular directions in the surface are everywhere equal and opposite. Elements of surface are thus saddle-shaped and can be joined smoothly to give periodic structures. The surface's mean curvature is zero and its gaussian curvature $K_1 K_2$ is non-positive. If the pressures on the two sides of the soap film are unequal, the surface becomes one of constant mean curvature.

The main physical significance is that, if we are considering a soap film, the pressure difference between the two sides is proportional to the mean curvature and to the surface tension. When an inextensible surface is deformed, the gaussian curvature is preserved. In the simplest case, a plane sheet, having zero gaussian curvature everywhere, can be deformed (developed) into a cylinder or a cone, which are also surfaces of zero gaussian curvature everywhere, but not into the surface of a sphere.

Surfaces like the D-surface have become popular for describing structures on various scales, from the cuprite structure to the children's climbing labyrinths built by Peter Pierce. The equivalent pieces of surface may be curved as in soap films or may be represented as flat panels. Statistical and disordered variants can also be found. Some surfaces divide space into two congruent regions and some into different regions. The intriguing G-surface divides space into two regions that are mirror images of each other. Figure 2 shows Schwarz's diagram of the unit tetragonal cell of the T-surface to which the interface of Thomas *et al.* like 'Scherk's minimal surface' (their Fig. 2 on page 600) is related. The T-surface is obtained by repetition of the interface which they observe.

Several groups are exploring the geometry of such surfaces, some new and some rescued from mathematical obscurity, and numerous physical applications have been made to silicates, liquid crystals, metals, polymers and biological materials. Thomas and co-workers have found an interesting realization of this diamond structure as the interface between the components in a two-component polymer. It has a periodicity of some 1,000 Å. Their structure is extraordinary in that it is composed of three separate regions, say A, B and C, each continuous in three dimensions. A is one of the diamond networks with the links inflated; B is the other interlocked diamond network. A and B are made of polyisoprene. C is a sheet which everywhere separates the A and B networks and it is made of polystyrene. Thus C has interfaces with A and B, A with C, B with C but not A with B.

Region C could be regarded as a thickened D-surface (of polystyrene) in a polyisoprene medium. The same D-surface has also been found by W. Longley and T. J. McIntosh (*Nature* 303, 612–614; 1983) in a mixture of glyceryl monooleate and water.

Transitions between continuous and disconnected interfaces between two phases can be envisaged. For example, if one phase consists of small spheres of one component embedded in an equal volume of the other component, the spheres could touch to join up, giving a

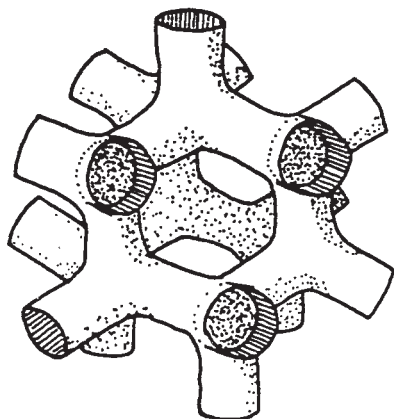


Fig. 1 The diamond network with the bonds represented by tubes. (From Williams, R. *The Geometrical Foundation of Natural Structure*, Dover, New York, 1972.)

uous honeycomb interface between the two phases.

A similar mechanism could be involved in the transformation of layered materials such as kaolin to three-dimensionally periodic networks such as ultramarine. If the sandwich layers are closer than some critical distance, unstable tubes like those

shown in Thomas and co-workers' Fig. 3a on page 600 of this issue could grow between the surfaces to minimize the interfacial area. □

Alan L. Mackay is in the Department of Crystallography, Birkbeck College, London WC1E 7HX, UK.

Population dynamics

More time means more variation

John H. Lawton

UNDERSTANDING patterns in the amplitude and frequency of population fluctuations is one of the central problems in ecology¹⁻³. Why do some populations fluctuate very little, whereas others vary wildly? Why do a minority of populations vary cyclically when the majority show no clear periodicity, and so on? Implicit in these questions is the assumption that the variation itself is easy to measure and hence that characteristic levels of population variation can be defined for various taxa³⁻⁵. However, as Pimm and Redfearn show⁶ on pages 613-614 of this issue, quantifying the variation itself is not a straightforward problem. The longer you observe a population, the more variable it seems to be. This simple fact has important theoretical and practical implications.

Pimm and Redfearn examined data from a large number of terrestrial animal populations (4 species of insects and 26 of birds and mammals) from censuses taken at least annually for over 50 years, plus information on 42 species of farmland birds and 32 species of woodland birds spanning 24 years, gathered by the British Trust for Ornithology. They measured population variation as the standard deviation of the logarithms of annual densities^{3,7}, and applied this measure to increasingly long segments of the counts for each population (2, 4, 8 years and so on). In most cases, more time meant more variation, right up to the maximum length of the time series. Very roughly, populations observed over 20 or 30 years seem to be twice as variable as populations studied for just 2 or 3 years (see figure).

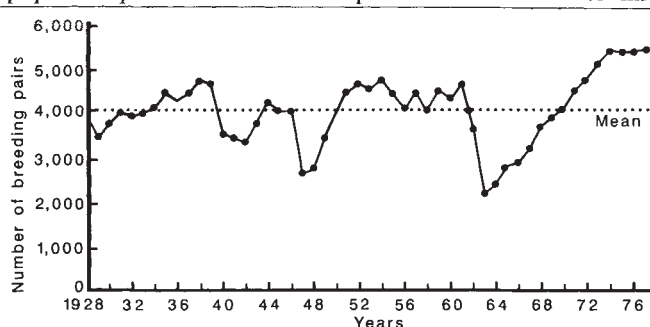
As the authors point out, an explanation for this observation almost certainly resides in the nature of environmental variation. Environmental disturbances, such as hard winters, fires and unusually good summers, that buffet or boost populations irrespective of density — the so-called density-independent processes¹⁻³ — are a major cause of population fluctua-

tions. We are all familiar with the idea that the further two places are apart in space, the less similar will be their environments: Chicago and Calgary have different climates but they are more similar to one another than either is to Calcutta. It is perhaps not so immediately obvious that the same is true at one place through time⁸. Mathematically, this is the same as saying that environmental variation at one place has a reddened spectrum. The term

Redfearn's observations⁶ have several important theoretical and practical implications. For example, work on comparative population variability^{4,5,9,10} will be much more difficult to interpret if it mixes studies that differ in duration. Second, do all populations continue to get more variable as time goes on or does variation level out? A truly reddened spectrum would imply that variation increases inexorably over time. Yet there are signs in Pimm and Redfearn's data that variation rises asymptotically to a maximum, at least in some populations. One possibility is that intrinsic stabilizing mechanisms (density-dependent processes) can constrain population variation, at least for a time, before eventually being overwhelmed. The fate of all populations is ultimate extinction.

The implications for conservation biology are profound. An increasing proportion of the exciting and beautiful animals in the world are confined to fragments of their former range in reserves and national parks and we need urgently to know whether these small, remnant populations are viable in the long run. A key parameter in increasingly sophisticated models used to predict minimum viable population sizes¹¹ is the variance r , that is, in the intrinsic rate of increase. Existing models assume that environmental variation driving variation in r takes particular, average values; but, as we now know, it would be much more realistic to assume that environmental variation increases over time. In other words, minimum viable populations calculated using a reddened spectrum for environmental noise must inevitably be bigger than current estimates based on constant variation. For many species, the long-term

survival of small, isolated populations looks bleak indeed. □



Number of breeding pairs of grey herons *Ardea cinerea* in England and Wales, 1928-77, based on annual censuses carried out by the British Trust for Ornithology¹². The study constitutes the longest continuous series of population figures available for any European breeding bird. The population has remained relatively stable over this 50 year period (with an average of about 4,000 pairs). Nevertheless, it is clear that the amplitude of population fluctuations (revealed by absolute maximum and minimum counts) has tended to increase with time, as shown by Pimm and Redfearn⁶ for a large number of other animal populations. (Modified from ref. 12.)

is an analogy with visible spectra in which long wavelengths are red and short ones are blue. If environmental variation, rainfall or temperature, for example, is modelled using a series of superimposed sine waves, it turns out that amplitudes increase towards the longer frequencies, that is, 'red' wavelengths predominate. More formally, when a spectral analysis is undertaken of environmental data, the amplitude of long wavelengths is consistently greater than that of short wavelengths, yielding an approximately straight-line relationship between log variance and log frequency⁸. So the longer we observe particular populations, the more likely we are to see increasingly extreme excursions in density, driven by increasingly extreme density-independent perturbations.

Intrinsic interest aside, Pimm and

1. Krebs, C.J. *Ecology. The Experimental Analysis of Distribution and Abundance* (Harper & Row, New York, 1985).
2. May, R.M. (ed.) *Theoretical Ecology. Principles and Applications* (Blackwell, Oxford, 1981).
3. Williamson, M. *The Analysis of Biological Populations* (Arnold, London, 1972).
4. Hassell, M.P., Lawton, J.H. & May, R.M. *J. anim. Ecol.* **45**, 471-486 (1976).
5. Connell, J.H. & Sousa, W.P. *Am. Nat.* **121**, 789-824 (1983).
6. Pimm, S.L. & Redfearn, A. *Nature* **334**, 613-614 (1988).
7. Williamson, M. *Ecol. Ent.* **9**, 239-241 (1984).
8. Williamson, M. in *Colonization, Succession and Stability* (eds Gray, A.J., Crawley, M.J. & Edwards, P.J.) 353-371 (Blackwell, Oxford, 1987).
9. Gaston, K.J. & Lawton, J.H. *Nature* **331**, 709-712 (1988).
10. Redfearn, A. & Pimm, S.L. *Ecol. Monogr.* **58**, 39-55 (1988).
11. Belovsky, G.E. in *Viable Populations for Conservation* (ed. Soulé, M.E.) 35-57 (Cambridge University Press, 1987).
12. Reynolds, C.M. *Bird Study* **26**, 7-12 (1979).

John H. Lawton is a professor in the Department of Biology, University of York, York YO1 5DD, UK.

Origins of life

Hydrothermal vents too hot?

Gerald Joyce

DID life on Earth originate in the outflow from a submarine hydrothermal vent? After careful examination of the relevant prebiotic chemistry, Miller and Bada, on page 609 of this issue¹, conclude that this is exceedingly unlikely. The hydrothermal vents are certainly a remarkable environment and are thought to have changed very little over the past 4×10^9 years of Earth history. They offer a carbon source which is exposed to high temperature (above 300 °C) under strongly reducing conditions. However, Miller and Bada consider such conditions to be unfavourable for the synthesis of biological polymers in aqueous solution.

The discovery of large-scale hydrothermal activity at submarine ridge crests is one of the most significant findings in oceanography. A volume of water equivalent to the entire ocean passes through the ridge lavas every 8–10 million years (Myr), profoundly affecting the chemical balance of the ocean and of the Earth as a whole². Because this is a phenomenon that has been operating continuously since before life began, it is reasonable to ask whether it had some influence on the chemical events that led to life's origins. Undoubtedly there were indirect effects resulting from the influence on the global environment. It has been proposed that a direct connection existed as well: that life began in the outflow tract of a hydrothermal-vent system³.

There are several features of the environment of a hydrothermal vent that make it an attractive candidate for the origin of life. The main biogenic elements were present in the form of H_2 , N_2 , H_2S , CO , CO_2 and possibly CH_4 ; several metal ions were available, including Fe^{2+} and Mn^{2+} ; the hot magma chamber provided a continuous source of thermal energy; and the entire system was shielded by the ocean from the destructive effects of ultraviolet irradiation. It has recently been pointed out that a submarine environment is also relatively protected from the deleterious effects of meteoric impacts, which were frequent before 4,000 Myr ago⁴. Finally, there is the intuitive notion, which dates back to the pre-socratic philosophers, that life is a geophysical phenomenon. Archelaus proposed in about 450 BC that "when the Earth was first being warmed,

in the lower part where the warm and cold were mingled together, many living creatures appeared . . . deriving their sustenance from the slime"⁵. Setting intuition aside, one must ask whether hydrothermal-vent conditions are consistent with the chemistry required to synthesize sugars, proteins and nucleic acids.

Miller and Bada¹ note that the high-temperature synthesis of amino acids is a very inefficient reaction in an aqueous environment. Whatever amino acids were

Subsequent investigation showed that the evidence for bacterial growth was likely to be the result of artefacts introduced during sample processing⁷. There is still no solid evidence for organisms growing at temperatures above 105 °C. Biology has apparently been unable to overcome the physical limitations imposed by extreme temperature conditions — it would be surprising if the first living organisms had managed to do so from the start.

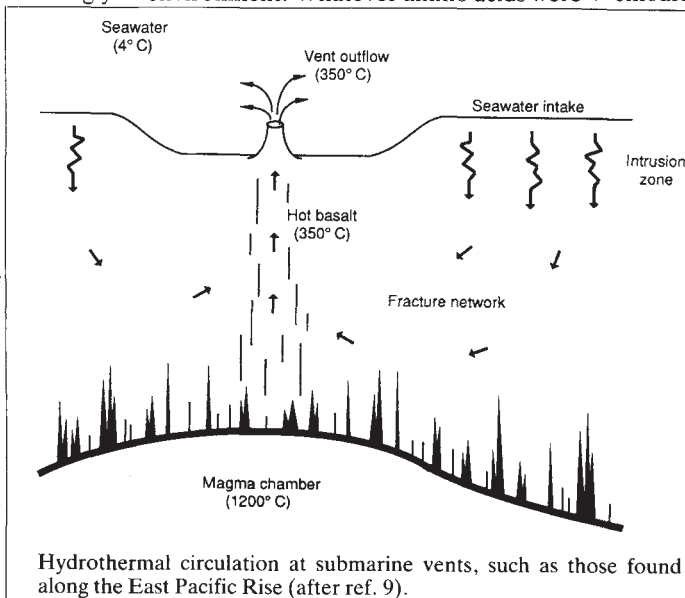
This, of course, raises the question of what is meant by 'the start'. "Most biologists would agree that every living system must at the very least be capable of both replication and evolution."⁸ The first living system, therefore, was an entity that was capable of evolving but had not yet embarked on a particular evolutionary pathway. Without the benefit of evolutionary improvement, such an entity must have been biochemically inept in the extreme. It is almost inconceivable that it could have been anything other than a heterotroph, deriving structural materials and energy from compounds that already existed in the environment. Theories suggesting that the first living organisms arose in hydrothermal vents and obtained their energy by carrying out anaerobic fermentation^{3,9} must address the problem of how primitive metabolic pathways arose spontaneously in the prebiotic environment.

Is there any hope, then, for the hydrothermal-vent hypothesis for the origins of life?

The problems of conducting useful chemical syntheses in the high-temperature outflow tract seem to be insurmountable. Perhaps life originated in low-temperature vents, such as those at the Galapagos rift or in the intrusion zone surrounding the site of hydrothermal discharge⁹. One would then have to postulate an energy source other than contact with the underlying hot basalt, which Miller and Bada show to be incompatible with the synthesis of biological polymers. The further one backs away from the hot magma chamber, the more the hydrothermal-vent hypothesis begins to sound like other theories that do not invoke a direct connection between the vents and the origins of life. □

1. Miller, S.L. & Bada, J.L. *Nature* **334**, 609–611 (1988).
2. Edmond, J.M. *et al.* *Nature* **297**, 187–191 (1982).
3. Corliss, J.B. *et al.* *Oceanologica Acta Proc. 26th Int. Geol. Conf.* 59–69 (1981).
4. Maher, K.A. & Stevenson, D.J. *Nature* **331**, 612–614 (1988).
5. Abel, E.L. *Ancient Views on the Origins of Life*, 49 (Fairleigh Dickinson University Press, Rutherford, New Jersey, 1973).
6. Baross, J.A. & Deming, J.W. *Nature* **303**, 423–426 (1983).
7. Trent, J.D. *et al.* *Nature* **307**, 737–740 (1984).
8. Weiner, A.M. in *Molecular Biology of the Gene* (eds Watson, J.D. *et al.*) 1104 (Benjamin/Cummings, 1987).
9. Baross, J.A. & Hoffman, S.E. *Origins of Life* **15**, 327 (1985).

Gerald Joyce is at the Salk Institute for Biological Studies, La Jolla, California 92037, USA.



produced would have decomposed within a few minutes under hydrothermal-vent conditions. The half life of sugars and of hydrogen cyanide would have been of the order of seconds, so it is difficult to imagine how mononucleotides could have formed. In an aqueous environment, peptide synthesis by thermal dehydration is out of the question. Peptide synthesis using prebiotic condensing agents such as hydrogen cyanide and cyanamide remains a possibility, although these agents are highly susceptible to hydrolysis at 200–300 °C and 200–400 atmospheres (atm) pressure. One would have to propose a mechanism to ensure their rapid use in driving the synthesis of more stable compounds. The lability of peptide bonds and the extreme lability of internucleotide bonds under hydrothermal-vent conditions poses yet another difficult problem.

In 1983, Baross and Deming⁶ described a community of extreme thermophiles that were cultured from the 350 °C waters of the hydrothermal vents along the East Pacific Rise. They claimed that these 'black smoker' bacteria could be grown in the laboratory at 250 °C and 265 atm.

Atmospheric physics

Climate forcing by warm seas

Michael Davey

VARIATIONS with a period of 30–50 days are an important feature of the Earth's climate, being implicated in monsoons and El Niño events. Evidence of the oscillations is found in zonal wind patterns and even in millisecond variations in the length of day, measured by lunar laser ranging. In a new examination of extensive data sets¹, Krishnamurti *et al.* show that fluctuations in sea surface temperature are correlated with the wind variations. Also, their calculations, which specifically exclude any interaction between the troposphere and the stratosphere, confirm that there is a strong air–sea coupling that can maintain the 30–50-day waves by the heat flux from the oceans to the troposphere.

Large-scale (encompassing the Equator) low-frequency oscillations in the atmosphere² appear throughout the year as waves propagating eastwards with fluctuating amplitude and period, which are generally strongest over the Indian and western Pacific oceans. They are closely associated with convection and rainfall and are powered by the strong release of latent heat in these warm regions³. Further evidence of the waves comes from satellite measurements of outgoing long-wave radiation (OLR), a measure of deep convection in the tropical troposphere. The dense global nature of the OLR data allows detailed studies of the fluctuations to be made⁴: the 30–50-day waves account for up to 25 per cent of the total OLR variance.

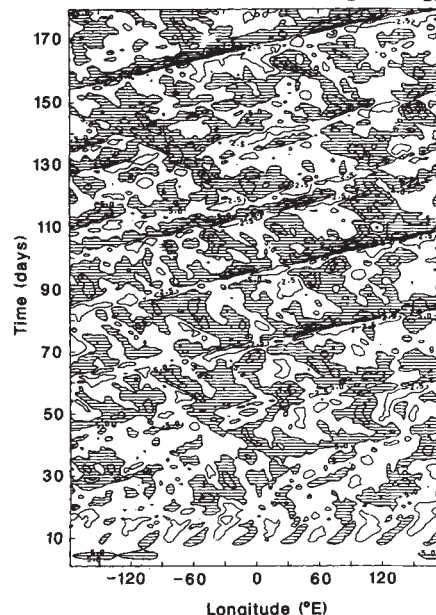
Apart from their intrinsic interest as atmospheric phenomena which require explanation, the oscillations are important because they are connected with the onset, active (wet) and break periods of the Indian⁵ and Australian monsoons. As the low surface-pressure component of the wave passes the Indian Ocean region, low-level convergence is enhanced and convection with associated rainfall can occur more easily. The structure in the Indian region has been extensively investigated, and 30–50-day waves moving northward from the Equator towards the Himalayas have also been identified. It is not known whether this link with the monsoon can be used reliably for long-range predictions.

There could also be a connection with longer timescales, as there are indications that the intensity of the 30–50-day waves is related to interannual El Niño–Southern Oscillation (ENSO) events⁶. Such waves could even act as the trigger for some ENSO events, by their association with wind and sea-surface-temperature (SST) anomalies that can be amplified slowly

by air–sea interactions under the right conditions.

Despite their importance, the cause of these waves is not yet fully understood. Their eastward propagation, equatorial prominence and baroclinic structure are consistent with the properties of a thermally forced equatorial Kelvin (transverse) wave. But the period is a puzzle because it fits the timescale of neither free atmospheric waves nor of forcing such as the seasonal cycle.

A plausible explanation is that the speed of the wave is intrinsically connected with its latent heating energy



Time-longitude map of convective rain (mm per day) averaged between 20° N and 20° S in an 'aquaplanet' experiment. The numerical model is essentially a weather-prediction model with land processes removed and SST fixed and zonally uniform. Waves propagating eastwards arise spontaneously, being less clear in the patchy smaller-scale rainfall than in the smoother velocity potential.

source: it must propagate at a speed that allows energy to be fed into the wave to maintain it against the dissipation of surface friction and radiative cooling. Even within this framework several mechanisms have been proposed. In the 'wave-CISK' theory⁷ cited in the new paper by Krishnamurti *et al.*¹, latent heating (by energy released as evaporated sea water recondenses) increases the buoyancy of a rising parcel of air, thus reducing the gravitational restoring force and considerably decreasing the speed of a gravity wave. Alternatively, the latent heating could be sensitive to surface evaporation rates which in turn depend on winds asso-

ciated with the wave; this feedback also yields oscillations of similar, intraseasonal period⁸. The reason for the unusual wave period could even be external rather than internal, as cold surges from higher latitudes are related to the intensification of equatorial convection⁹. Krishnamurti *et al.*¹ also list instabilities of zonal winds among the mechanisms.

Analyses of existing observations are important to help elucidate the various possibilities. Krishnamurti *et al.*¹ use atmospheric data sets acquired during the 1979 monsoon experiment and the 1978–79 global experiment, together with SST measurements, and find a signal of 30–50 days in the atmosphere and ocean. They also calculate the sensible and latent surface-heat fluxes (carried by convection and evaporation, respectively), concentrating on sites in the central Arabian Sea, the Bay of Bengal and the equatorial western Pacific where the signals are prominent. Fluctuations in the 30–50-day range account for about 20 per cent of the total latent heat flux of around 200 W m⁻² and for about 10 per cent of the much smaller sensible heat flux.

The heat fluxes depend on surface winds, humidity, and air and sea temperatures. To assess the importance of each variable, Krishnamurti *et al.* suppress fluctuations in one or more and recalculate the fluxes. For 30–50-day periods, the suppression of wind variability has the largest effect on the latent heat flux. SST also has a large effect, but air temperature and humidity have little. The results demonstrate that there is substantial coupling of the ocean to the atmosphere for intraseasonal timescales, but cannot reveal any mechanism. Through their effect on local latent-heat fluxes, the variations in SST could contribute significantly to the generation and maintenance of the atmospheric waves, but the SST changes could also be driven by the atmosphere without substantial feedback.

An extra incentive for determining the causes of these waves is that numerical weather-prediction models generally do not simulate them well, instead producing waves with a shorter 20–30-day period. The period is related to the way deep convection is modelled, so a proper understanding of 30–50-day waves should be reflected in improved representations of convection and rainfall. The results presented by Krishnamurti *et al.*¹ also show

1. Krishnamurti, T.N., Oosterhof, D.K. & Mehta, A.V. *J. Atmos. Sci.* **45**, 1304–1322 (1988).
2. Madden, R.A. & Julian, P.R. *J. Atmos. Sci.* **28**, 707–708 (1971).
3. Madden, R.A. *J. Atmos. Sci.* **43**, 3138–3158 (1986).
4. Weickmann, K.M., Lussky, G.R. & Kutzbach, J.E. *Mon. Weath. Rev.* **113**, 941–961 (1985).
5. Krishnamurti, T.N. *et al. J. Atmos. Sci.* **42**, 364–375 (1985).
6. Lau, K.-M. & Chan, P.H. *J. Atmos. Sci.* **45**, 506–521 (1988).
7. Lau, K.-M. & Peng, L. *J. Atmos. Sci.* **44**, 950–972 (1987).
8. Neelin, J.D., Held, I.M. & Cook, K.H. *J. Atmos. Sci.* **44**, 2341–2348 (1987).
9. Wang, X.-L. & Murakami, T. *J. met. Soc. Jap.* **65**, 727 (1987).
10. Hayashi, Y.-Y. & Sumi, A. *J. met. Soc. Jap.* **64**, 451 (1986).

that improved data on SST are needed.

Despite their defects, numerical models have the advantage that the physics can be altered in controlled experiments to study various mechanisms. An interesting result is that the oscillations appear even when the Earth is completely covered with ocean, the SST constant and the seasonal

cycle frozen¹⁰ (see figure). Thus the omitted processes are not fundamental to the existence of intraseasonal atmospheric waves, although they could have important modulating effects. □

Michael Davey is at the Meteorological Research Unit, Hooke Institute, Oxford OX1 3PU, UK.

Palaeontology

Bringing up baby

Michael J. Benton

THE assumption that small dinosaurs are always young dinosaurs is challenged by recent work on fossil reptiles from the Early Permian of North America¹. In a study of 392 limb bones from two pelycosaur genera *Ophiacodon* and *Dimetrodon*, Brinkman shows that bone shape, conditioned as it is by developmental stage, tells us more about growth than does gross size alone.

The data allow Brinkman to define five growth stages in the humerus of *Ophiacodon* (Fig. 1), reflecting a change in shape in the ends of the bones from concave to convex as juvenile cartilage is replaced by bone, accompanied by the progressive and distinctive development of specialized processes that provide either muscle attachments or joint surfaces with other arm bones. But there is little correlation between Brinkman's five growth stages and bone size (Fig. 2), suggesting that size and age are not easy to link. Clearly the stage I humeri were smaller than those at stage V but the overlap of adjacent growth-stage classes was very great.

Brinkman argues¹ that his growth stages are better correlates of age than would be any arbitrary set of size classes. This is because growth stage relates almost solely to age, whereas body size could be affected by factors such as sexual dimorphism, geographic variation, habitat differences or the unexpected

presence of more than one species or subspecies in the fossil sample.

The importance of shape change with growth does not diminish the usefulness of size comparisons in certain circumstances. Discoveries of juvenile skeletons in eggshells² provide independent evidence that smallness can imply youth, but such discoveries are extremely rare. More common are fossil reptiles so much smaller than the known adults that their age is fairly evident. For example, the discovery of a tiny skeleton of a marine nothosaur, just 51 mm long, in Middle Triassic rocks in Switzerland³ compares with adults of the same species ranging in length from 230–370 mm found in the same deposits.

Equations have been used to estimate the age of the tiny nothosaur³. The first equation is based on the size ratios of adults to hatchlings of 120 species of modern lizards and crocodilians⁴. Because of the similar cold-blooded physiologies of lizards and crocodilians, these reptiles grow at similar rates and it is possible to derive a simple equation relating age and size. A second equation is based on a study⁵ of the size ratios of adults to hatchlings of a broader cross-section of modern reptiles, including turtles and snakes. These equations can be used, tentatively, to calculate the expected hatchling size of a fossil reptile whose adult size is known.

Such an application is based on the assumption that the fossil reptile had similar growth rate and physiology to modern reptiles, and that enough is known of the population structure of the fossil species to identify the typical adult size.

The question of whether small dinosaurs are always juveniles or not has been considered by Callison and Quimby⁶. They use two main non-size-related criteria, the degree of ossification, that is, replacement of cartilage by bone, as Brinkman now also emphasizes, and the relative proportions of different elements of the skeleton. It is well known that the proportions of all vertebrates change during growth. A human baby, for example, has a relatively huge head, short legs and big hands when compared with an

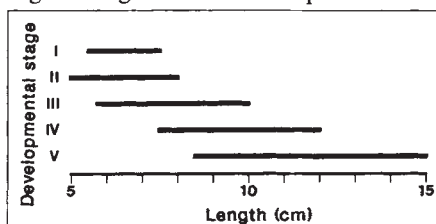


Fig. 2 The relationship between size and stage of development of a sample of 27 humeri from localities in west-central Texas. Particular growth stages correspond to a broad spectrum of body sizes. (From ref. 1.)

adult. Callison and Quimby compare a number of small bipedal dinosaurs, some no larger than chickens, with a broad range of modern running birds such as chickens, turkeys and ostriches. A baby ostrich is the same size as an adult chicken but it has huge knees, in anticipation of the adult requirements, similar to the large feet of puppies.

A graph of femur (thigh bone) length against knee width (the width of the lower end of the femur) shows that the relative proportions correlate closely with age in modern birds. Bipedal dinosaurs have very similar femurs, so that it is probably legitimate to compare the two groups. The chicken-sized dinosaurs are interpreted as adults on the basis of their slim-line knees.

The new studies of the growth of fossil reptiles show that size alone is not necessarily a good measure of age. Better measures may be based upon the relative proportions of certain skeletal elements and upon the degree of ossification of individual bones. □

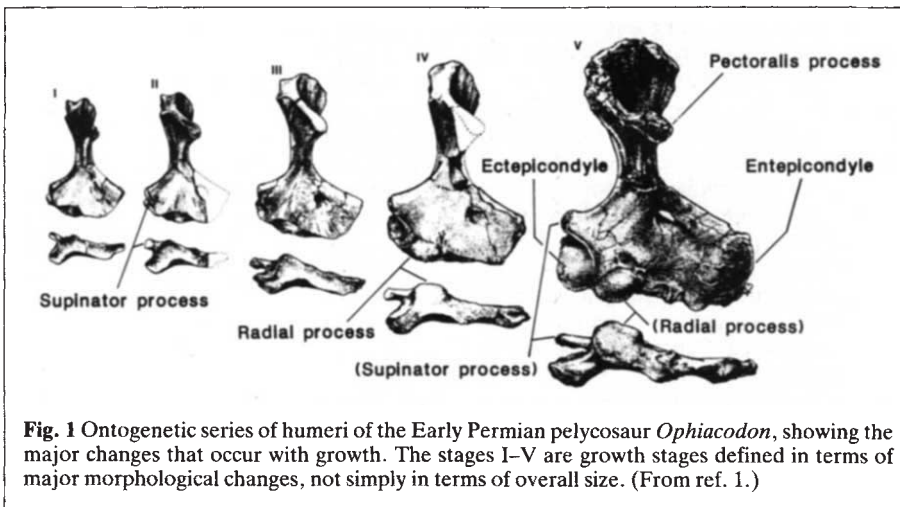


Fig. 1 Ontogenetic series of humeri of the Early Permian pelycosaur *Ophiacodon*, showing the major changes that occur with growth. The stages I–V are growth stages defined in terms of major morphological changes, not simply in terms of overall size. (From ref. 1.)

1. Brinkman, D. J. *vertebr. Paleont.* **8**, 172–180 (1988).
2. Horner, J.R. & Weishampel, D.B. *Nature* **332**, 256–257 (1988).
3. Sander, P.M. *Science* **239**, 780–783 (1988).
4. Currie, P.J. & Carroll, R.L. *vertebr. Paleont.* **4**, 68–84 (1984).
5. Andrews, R.M. in *Biology of the Reptilia* (eds Gans, C. & Pough, F.H.) 280–282 (Academic, London, 1982).
6. Callison, G. & Quimby, H.M. *vertebr. Paleont.* **3**, 200–209 (1984).

Michael J. Benton is in the Department of Geology, The Queen's University, Belfast BT7 1NN, Northern Ireland.

Tubulin synthesis

Controlling mRNA lifespan

Tim Hunt

MESSSENGER RNA in animal cells is generally very stable, with half lives of the order of several hours being the norm. Thus, the level of ribonuclease in cytoplasm must be extremely low, with what little there is probably held in check by RNasin, the RNase inhibitor which is found in liver, reticulocytes and placenta (probably all cells, in fact). This being so, it is the scattered cases of unstable mRNA that require explanation. The case of β -tubulin mRNA is a well-known example. It has been known for many years that tubulin synthesis is shut off by drugs that disassemble microtubules, and is enhanced by microtubule stabilizers like taxol¹. Neither transcriptional nor translational control plays a significant part in the regulatory circuit; colchicine causes tubulin mRNA to become unstable and, conversely, taxol stabilizes it. The effect is specific for tubulin mRNA, and all the evidence suggests that unpolymerized $\alpha\beta$ dimers regulate their own synthesis by a feedback mechanism acting on their mRNA. Studies of β -tubulin mRNA in Cleveland's laboratory at Johns Hopkins University pinned down the regulatory sequences to the 5' end of the gene². This

in itself is unusual; destabilizing sequences usually lie in the 3' untranslated regions (well documented examples are found in histones³, GM-CSF (granulocyte-monocyte colony stimulating factor)⁴, *c-fos*^{5,6} and the transferrin receptor^{7,8}). But, even more surprisingly, the latest work from Cleveland's group, reported on page 580 of this issue⁹, shows that the instability of the tubulin mRNA depends on recognition of the first four amino acids of β -tubulin, Met-Arg-Glu-Ile (MREI), as they emerge from the ribosome, and not on the sequence or structure of the mRNA as such.

The evidence for this view is comprehensive and elegant. Alterations to the sequence of the mRNA permit normal regulation as long as the N-terminal amino-acid sequence stays the same, whereas single nucleotide changes that alter the protein sequence stop the regulation. Regulation works only when the sequence is translated in the correct frame, and the peptide sequence must be located at the N terminus of the nascent chain. Finally, the sequence must be followed by a reasonable length of polypeptide, long enough for it to emerge from the ribosome. If the protein containing the sequence is too short, or if peptides are released prematurely by puromycin, the mRNA is stabilized. These observations suggest a model in which something, presumably a nuclease or a nuclease activator, binds to ribosomes via the N terminus of nascent β -tubulin when the concentration of $\alpha\beta$ -tubulin dimers gets too high. Perhaps free tubulin dimers bind to the sequence MREI, carrying a nuclease with them, but how such a mechanism might work is unclear.

How do these findings fit with what is known of other examples of regulated mRNA stability? The mechanism of mRNA turnover is still poorly understood, but several recent observations begin to suggest a fairly solid outline. The first significant result is that intermediates in the destruction of the mRNA are rarely, if ever, seen. This is most easily understood if the cell contains exonucleases that rapidly degrade the fragments of mRNA produced by the first (rare) endonucleolytic cut. Why do these enzyme(s) not attack intact mRNA? Almost certainly it is protected at both ends, at the 5' end by the 'cap' structure and at the 3' end by the poly(A) tail. Thus, a single endonucleolytic clip renders it vulnerable to the scavenging exonucleases. Ross and his colleagues have now identified a 3' $\rightarrow$ 5' exonuclease which is

probably responsible for part of this process, and show that histone mRNA is degraded 3' $\rightarrow$ 5' *in vitro*¹⁰. Histone mRNA lacks a poly(A) tail and should therefore be highly vulnerable to destruction by this activity; however, it contains a structure at the 3' end which provides partial and regulated protection³. The story is complicated, with echoes of the tubulin story in that destruction of the mRNA requires active translation of the mRNA and the accumulation of free (that is, not bound to DNA) histones in the cytoplasm¹¹. But there is as yet no evidence for any involvement of the nascent chain in this case; although like Cleveland, Marzluff and his colleagues propose¹¹ that the translating ribosomes carry a nuclease with them.

Two other recent papers provide support for the basic idea that the ends of mRNAs are sensitive targets for destruction. Thus, Harland and Misher made a circular (endless) mRNA and injected it into amphibian (*Xenopus*) embryos¹². In agreement with the model, such RNA molecules are exceptionally stable. These circles should be a useful test bed for potential destabilizing sequences and the factors that attack them. The only flaw in this idea is that circular mRNAs cannot be read by ribosomes¹³, so the stability of these structures might arise as much from their failure to be translated as from their endlessness. Again using *Xenopus*, Hyman and Wormington¹⁴ report that the translational inactivation and disappearance of the mRNAs for ribosomal proteins that accompanies oocyte maturation is preceded by removal of their poly(A) tails. Something of the same sort happens to *c-myc* and *c-fos* mRNA, where degradation is apparently preceded by shortening of the poly(A) tail^{15,16}. This is a specific process which requires sequences in the 3' untranslated region of the mRNAs; stable mRNAs like globin do not show poly(A) shortening. Other mRNAs in *Xenopus* oocytes suffer precisely the opposite fate, getting longer poly(A) tails at the same time that the ribosomal

1. Ben Ze'ev, A., Farmer, S.R. & Penman, S. *Cell* **17**, 319-325 (1979).
2. Gay, D.A., Yen, T.J., Lau, J.T.Y. & Cleveland, D.W. *Cell* **50**, 671-679 (1987).
3. Marzluff, W.F. & Pandey, N.B. *Trends biochem. Sci.* **13**, 49-52 (1988).
4. Shaw, G. & Kamen, R. *Cell* **46**, 659-667 (1986).
5. Treisman, R. *Cell* **42**, 889-902 (1985).
6. Fort, P. *et al. Nucleic Acids Res.* **15**, 5657-5667 (1987).
7. Casey, J.L. *et al. Science* **240**, 924-928 (1988).
8. Müllner, E.W. & Kühn, L.C. *Cell* **53**, 815-825 (1988).
9. Yen, T.J., Machlin, P.S. & Cleveland, D.W. *Nature* **334**, 580-585 (1988).
10. Ross, J., Kobs, G., Brewer, G. & Pelz, S.W. *J. biol. Chem.* **262**, 9374-9381 (1987).
11. Graves, R.A., Pandey, N.B., Chodchoy, N. & Marzluff, W.F. *Cell* **48**, 615-626 (1987).
12. Harland, R. & Misher, L. *Development* **102**, 837-852 (1988).
13. Kozak, M. *Nature* **280**, 82-85 (1979).
14. Hyman, L.E. & Wormington, W.M. *Genes Devel.* **2**, 598-605 (1988).
15. Brewer, G. & Ross, J. *Molec. cell. Biol.* **8**, 1697-1708 (1988).
16. Dworkin, M.B., Shrutkowski, A. & Dworkin-Rastl, E. *Proc. natn Acad. Sci. U.S.A.* **82**, 7636-7640 (1985).
17. Palatnik, C.M., Wilkins, C. & Jacobson, A. *Cell* **36**, 1017-1025 (1984).
18. Sheiness, D., Puckett, L. & Darnell, J.E. *Proc. natn Acad. Sci. U.S.A.* **80**, 1077-1081 (1975).



100 years ago

I HAVE read with extreme interest the abstract of a paper by Prof. Ewart, on the "Structure and Development of the Electric Organ of *Raia radiata*." Organic nature is full of organs which are either wholly or partly functionless. The Darwinian philosophy almost invariably explains them as structures which must have once been useful, and have become functionless by atrophy or disuse.

This is a natural and necessary consequence of the doctrine which ascribes all organic structures to utility as a physical cause. Utility as a mental purpose is kept out of sight.

I have always thought that if the doctrine of development be true, functionless organs must be the germs of potential use, and not necessarily at all the remains of past actual use. Here we have a case in which a distinguished physiologist detects, or thinks he can detect, an organ in process of being built up for the discharge of a very definite and peculiar function.

This fact does not tell against development or evolution. But it does tell, and tells fatally, against the element of fortuity, which is inseparable from the idea of "natural selection."

ARGYLL

FROM *Nature* **38**, 341; 9 August 1888.

protein mRNAs are losing theirs. Many questions are outstanding in this system, in which translational activation and repression take place at the same time as programmed destruction¹⁶. There is some evidence that deadenylation reduces the affinity of the mRNAs for ribosomes¹⁷, but it has never been clear how changes at the 3' end of mRNA can influence ribosomes entering at the 5' end. Nevertheless, these observations all support the idea that one of the most important roles of the poly(A) tail is to protect mRNA against degradation, as originally proposed by Darnell and colleagues¹⁸. There is an excellent brief review of the role of poly(A) in determining mRNA life expectancy in the paper by Brewer and Ross¹⁵.

If it is true that mRNA is really very stable stuff, protected against attack from exonuclease by special structures at either end, the next advances should come from understanding the highly specific endonucleases that create suitable ends for the exonucleases to act upon, and the role that ribosomes seem to play in determining where and when they act. The β -tubulin system seems to offer a favourable opportunity for further insights, especially if the effects can be dissected *in vitro*. No doubt affinity columns bearing MREI are already sprouting in the cold rooms of Johns Hopkins. □

Tim Hunt is in the Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK.

Solid-state physics

A liquid of fractional charges

Gerhard Fasol

THE quantum Hall effect takes two distinct forms with very different properties, according to our current understanding. The first, the integral quantum Hall effect, has found practical application as a resistance standard as well as giving an improved value for the fine-structure constant α , a fundamental constant that describes the coupling of elementary particles to electromagnetic fields. The second, the fractional quantum Hall effect, manifests itself in high-quality samples with low charge-carrier concentrations in very high magnetic fields. It provides evidence of a totally new quantum liquid state of electrons which is predicted to show peculiar elementary excitations that behave like electrons, except that their electrical charge is less than that of the electron. Recent experiments by Clark *et al.*¹ substantially elucidate the properties of these peculiar quasiparticles, not only providing additional evidence of their existence but also measuring their charges directly.

A current I flowing through a conductor perpendicular to an applied magnetic field B becomes deflected sideways — the Hall effect. The resulting Hall voltage V_{xy} is measured in the direction perpendicular to both the flow of the current and the direction of the magnetic field. V_{xy} increases linearly with both I and B , and the quantity V_{xy}/I , which has the units of resistance, is termed the 'Hall' resistance although no dissipative heating is caused directly by it.

The integral quantum Hall effect (IQHE) was discovered in Nobel-prize-winning work² by von Klitzing and co-workers. These authors found that the Hall resistance of silicon-metal-oxide semiconductor field-effect transistors (MOSFETs) in a sufficiently high mag-

netic field rises through a series of quantum plateaux as the field is increased, not linearly as in the classical effect. The Hall resistivity at these plateaux is expressed to an accuracy of $1/10^7$ in terms of fundamental physical constants as $\rho_{xy} = h/ie^2$, where h is Planck's constant, e is the free electron charge and i is an integer. Moreover, this relationship is independent of both the sample geometry and materials used. The quantum Hall effect is now used to establish resistance standards and to determine accurately the fine structure constant α .

The integral effect occurs because the magnetic field breaks the continuous distribution of electron energies found in metals into a sequence of Landau levels, each of which can contain $N = eB/h$ electrons (Fig. 1). A system with n electrons has $\nu = n/N$ filled Landau levels. As B increases, the various Landau levels are brought into coincidence with the Fermi level — the typical energy of active

electrons — and more electrons can participate in the Hall effect, accounting for the stepwise increase in Hall resistance in the IQHE.

Störmer *et al.*³ fabricated modulation-doped ultra-high-mobility GaAs-(AlGa)As structures in which the electrons are separated from the charged donor impurities, thus dramatically reducing the normally predominant charged-impurity scattering. The Hall resistance ρ_{xy} in these samples, rises through plateaux at fractional filling factors ν , such as $1/3$, $2/3$, $2/5$ and $3/5$, in addition to the expected plateaux at integer filling values^{4,5} (Fig. 2). The Hall resistivity at these fractional plateaux is given by $\rho_{xy} = h/\nu e^2$. The parallel resistance ρ_{xx} vanishes at the magnetic-field values for which the plateaux in ρ_{xy} occur. The fractional quantized Hall effect (FQHE) has a cause that is totally different from that of the integral quantized Hall effect: a previously unknown quantum liquid state.

In a conventional isotropic conductor the conductivity σ_{xx} is just the inverse of the resistivity ρ_{xx} : $\sigma_{xx} = 1/\rho_{xx}$. In an anisotropic system (for which the diagonal resistivity $\rho_{xx} \neq 0$) this relationship is changed to $\sigma_{xx} = \rho_{xx}/(\rho_{xx}^2 + \rho_{xy}^2)$. Thus $\rho_{xx} = 0$ implies $\sigma_{xx} = 0$ as long as the condition $\rho_{xy} \neq 0$ is fulfilled.

The following example clarifies this seemingly paradoxical situation: electrons in mutually perpendicular magnetic and electric fields move in a cycloid orbit. They follow a track perpendicular both to the magnetic and to the electric field. There is no current in the direction of the electric field and no voltage in the direction of the current. Thus, the fact that ρ_{xx} vanishes at the magnetic field values of the fractional quantized Hall plateaux implies that the conductance σ_{xx} also vanishes there — with far-reaching consequences for our understanding of electrons in such extreme conditions.

Because electrons at the Fermi level are responsible for the parallel conductivity, the vanishing conductivity implies either

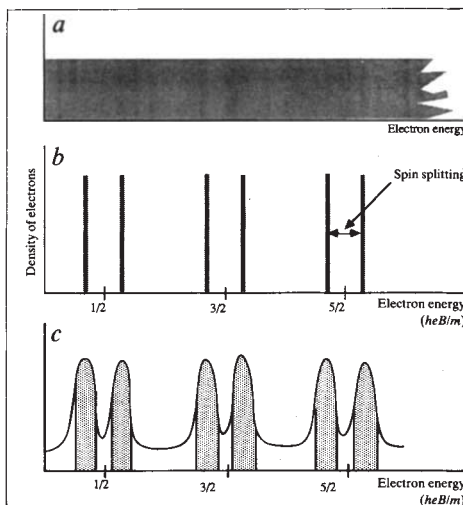


Fig. 1 Physics of the IQHE^{7,8}. Electrons in a two-dimensional system can have any energy up to about the 'Fermi' energy (a). Applying a magnetic field groups the allowed energies into quantized 'Landau' levels, each with up to $N = eB/h$ electrons (b). A system with n electrons has $\nu = n/N$ filled Landau levels: ν is the 'filling factor'. Imperfections and impurities broaden the Landau levels and make some electron states localized (non-conducting) (c), rendering the IQHE observable. Raising the applied field raises Landau levels past the Fermi energy. Localized states cannot contribute to the Hall effect, so the Hall resistance increases only as additional Landau levels rise above the Fermi energy, explaining the plateaux of ρ_{xy} . Similarly, the parallel resistance ρ_{xx} vanishes when the Fermi energy falls between Landau levels. (See also ref. 9).

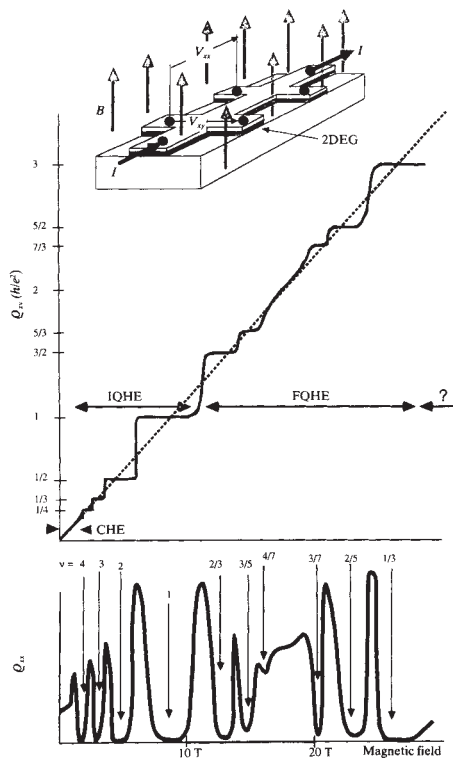


Fig. 2 The Hall and parallel resistances (Q_{xy} , top, and Q_{xx} , bottom) as measured in various experiments as a function of magnetic field. The inset shows the experimental arrangement (2DEG, two-dimensional electron gas; V , voltage). The classical effect, $Q_{xy} = B/en$, applies only at low fields. The integral quantum Hall effect (IQHE), with plateaux at $Q_{xy} = h/ie^2$, applies at intermediate fields and is observed in field-effect transistors with some imperfection and impurities. The fractional quantum Hall effect (FQHE), with plateaux at $Q_{xy} = h/ve^2$ ($v = p/q$) occurs at high fields in only the purest samples. The behaviour above 30 tesla is unknown and several groups are using pulsed magnets to achieve such extreme conditions.

that there are no electrons or that the electrons at the Fermi level are localized (non-conducting) for certain fractional filling values. The only way to explain such gaps in the spectrum of electrons is to treat the electrons collectively as a liquid with rather peculiar properties. Laughlin described many characteristics of this liquid by determining the nature of its wavefunction^{6,7}. A many-electron wavefunction is hard to visualize and it is easier to consider a single electron in the liquid. For this selected electron, Laughlin's wavefunction can be thought to have vortices where the probability of finding it is zero. The density of these 'forbidden spots' increases linearly with the applied magnetic field. Because electrons are governed by the Pauli exclusion principle, each electron in the system can be found only at the forbidden spots of all other electrons. It so happens that there is exactly one electron for each forbidden spot when the lowest Landau level is filled.

For partially filled Landau levels the system as a whole has a lower energy if forbidden spots merge to multiple

forbidden spots, meaning that the system tends towards this arrangement. For example, when the first Landau level is filled exactly up to one-third of its capacity, there will be exactly one electron in each threefold forbidden spot. Thus, this quantum liquid effectively locks the electrons into particular short-range patterns. Laughlin predicted that the excitations of this liquid state act as quasiparticles with a fractional charge, e/q , where q is the denominator of the appropriate fractional filling factor $v = p/q$. It is the condensation of the electrons into this new liquid state at fractional filling factors that is responsible for the fractional quantum Hall effect.

In the case of $v = 1/3$, the smallest possible excitation can be envisaged as the addition of a single vortex. As every threefold vortex is balanced by one electron of charge e , a single vortex corresponds to the change $+e/3$. There is an energy gap Δ necessary to excite a quasiparticle excitation of the quantum liquid. Because of the gap, no scattering and therefore no energy dissipation is possible at low temperatures so that Q_{xx} vanishes for this liquid state. Increasing the temperature generates quasiparticles. The parallel resistance at the various FQHE plateaux is related to temperature by the expression $Q_{xx} = Q_{xx}^c \exp(-\Delta/2kT)$ where Q_{xx}^c is the asymptotic resistivity.

By measuring Q_{xx} as a function of temperature for 13 different fractional states p/q , Clark *et al.*¹ could measure Q_{xx}^c and hence the charge of the quasiparticles. For a particular q , the values of Q_{xx}^c versus $1/T$ converge to the same value of Q_{xx}^c , regard-

less of the value of p . Using their measured values of Q_{xx}^c Clark *et al.* find that Q_{xx}^c scales as $1/q^2$, also independent of the value of p . Their experiment therefore gives strong evidence that the charge of the exotic FQHE quasiparticles is $e^* = e/q$, less than the fundamental charge of the electron, as predicted by Laughlin. The experiment supports the present understanding of the strange FQHE liquid state of electrons in the extreme quantum conditions found at high magnetic fields and low electron densities.

The extreme conditions necessary to observe the FQHE make it unlikely that any commercial devices based on the phenomenon will be available in the immediate future. But understanding the physics governing the behaviour of electrons under extreme conditions is essential for the development of ultrafast, submicrometre devices for computing and telecommunication systems. Elucidating the FQHE will also stimulate developments in magnetism and superconductivity, which are also governed primarily by interactions between electrons. □

1. Clark, R.G., Mallett, J.R., Haynes, S.R., Harris, J.J. & Foxon, C.T. *Phys. Rev. Lett.* **60**, 1747-1750 (1988).
2. von Klitzing, K., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45**, 494-497 (1980).
3. Störmer, H.L., Dingle, R., Gossard, A.C., Wiegmann, W. & Sturge, M.D. *Solid State Commun.* **29**, 705-709 (1979).
4. Tsui, D.C., Störmer, H.L. & Gossard, A.C. *Phys. Rev. Lett.* **48**, 1559-1562 (1982).
5. Störmer, H.L. *et al. Phys. Rev. Lett.* **50**, 1953-1956 (1983).
6. Laughlin, R.B. *Phys. Rev. Lett.* **50**, 1395-1398 (1983).
7. Laughlin, R.B. *Phys. Rev.* **B23**, 5632-5633 (1981).
8. Halperin, B.I. *Phys. Rev.* **B25**, 2185-2190 (1982).
9. Halperin, B.I. *Scient. Am.* **254**, 40-48 (April 1986).

Gerhard Fasol is in the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, CB3 0HE, UK.

Astrophysics

Reflected glory of active nuclei

Martin Ward

THE nuclei of some galaxies emit prodigious amounts of energy over a wide range of frequencies. How this energy is transmitted outward from the nucleus, and its influence on the host galaxy, is a topic of increasing interest to astronomers. Elsewhere in this issue (*Nature* **334**, 591-593; 1988) S. Alighieri and colleagues suggest that a beam of energy from the nucleus of the radio galaxy PKS2152-69 illuminates a bright cloud of gas about 8 kiloparsecs (2.4×10^{17} km) away.

Such energetic beams or cones of radiation are found in many active galaxies, and originate from regions close to the central black hole thought to power their active nuclei. The characteristics of the radiation we see depends on where we are relative to the axis of the beam (Fig. 1). In the case of PKS2152-69, the collimated emission is not directed towards us.

Last year several of the same researchers first identified a 'blob' of light about 8 kiloparsecs from the nucleus (Tadhunter, C.N. *et al. Nature* **325**, 504-507; 1987). Not only is it brighter and more blue than the surrounding region of the galaxy, but it has line spectra, characteristic of ionized atoms, superimposed on the continuum spectrum. They argued that a cluster of young stars within the blob could not produce the high ionization emission lines observed. Such stars would instead produce a lower ionization spectrum as seen in the Orion nebula in our Galaxy. So, an alternative source of energy is needed to explain the activity in the blob.

Crucially, Alighieri and colleagues show in their new paper in this issue that the continuum light from the blob is highly polarized. There are two ways in which this polarized light could arise; as synchrotron radiation from electrons spiralling in

a magnetic field; or from the scattering of a beam of photons by electrons or larger particles of dust. There are difficulties with both the synchrotron and electron-scattering processes, although they cannot be excluded without further observations. But the scattering of a beam of photons which hits a cloud containing gas mixed with dust is a possibility. We do not see the energetic photons directly because the angle between us and the beam is too great. These putative high-energy photons ionize the gas which exhibits the emission lines. Also, some photons are scattered into our direction of view, after they first collide with dust grains within the cloud.

This mechanism would explain both the high polarization of the continuum light and the fact that the polarization rises with increasing frequency. A crucial test of this model can be made once the much-delayed Hubble Space Telescope is launched. It will then be possible to measure the polarization curve into the ultraviolet regions, which is not observable from the Earth. The shape of the curve, in particular the frequency at which the polarization reaches a maximum, will

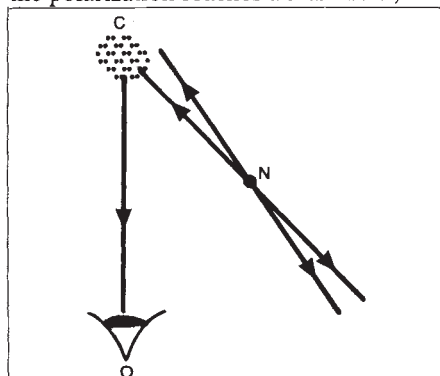


Fig. 1 The intense beam of radiation emitted from the nucleus (N) of an active galaxy can illuminate distant clouds (C) in the galaxy, as observed by Alighieri *et al.* (O, observer).

identify the process producing the polarization. If scattering is responsible, it will also show whether the scatterers are electrons of dust particles.

The authors suggest that their explanation can provide a link between different classes of active nucleus. The so-called blazars which have powerful continuum radiation but relatively weak emission lines are believed to have the axis of their collimated emission pointed nearly directly towards us. Objects like PKS2152-69 could be of the same type, but with a beam orientation that does not permit us to see down the collimated cone of photons and electrons. In the latter case the presence of an unseen blazar nucleus can be inferred indirectly, by observing the emission lines and scattered continuum that would be produced by collisions of the beam with material along its trajectory.

Studying the occurrence statistics of

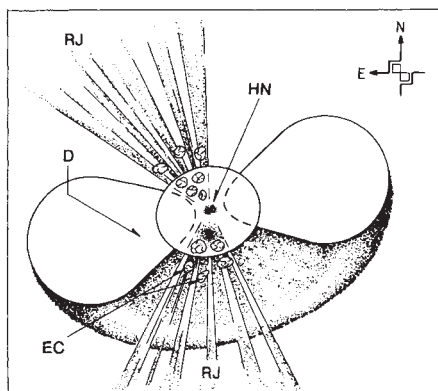


Fig. 2 Beams less collimated than those from blazars can be produced by funnelling along the poles of a dense disk (D) a few parsecs deep. The structure of NGC106B (from Chelli *et al. Astr. Astrophys.* 177, 51-61; 1987) shows a quasar-type nucleus hidden by the disk, and seen only via radiation reflected from clouds of electrons (EC) above and below the disk. RJ, radio jets; HN, hidden nucleus.

blazars that are not pointed towards us will be a non-trivial exercise. There could be many examples in which the beam escapes from the host galaxy without interacting with any material. Neither do we know what fraction of the total luminosity is scattered towards us, so any comparison of the properties of galaxies like PKS2152-69 with known blazars must be uncertain. Nevertheless, any such unification of the complex classification scheme for active nuclei is welcome.

A recent dramatic spectropolarimetric observation of the nearby galaxy NGC1068 (Antonucci, R.R.J. & Miller, J.S. *Astrophys. J.* 297, 621-632; 1985), which was long known to have relatively low-velocity-broadened optical emission lines, revealed a 'ghost' high-velocity component visible in polarized light. High-velocity gas is a property typical of quasar activity. In NGC1068, the quasar-type nucleus is completely hidden by a thick torus of molecular gas, and we can observe the nucleus only by the nuclear light which is scattered towards us by a cloud of electrons above the torus (Fig. 2). This observation opened up the whole question of how many powerful active nuclei might be hidden in this way.

The nuclear emission from PKS2152-69, in common with blazars and powerful radio galaxies, is probably highly collimated. Less collimated emission may result when an isotropically emitting source is funnelled out along the polar directions by either a thick accretion disk or on much larger scales, by a molecular disk. The scattering phenomenon can be thought of as a mirror, albeit a distorting mirror which reflects only a small fraction of the incident light. But it is a useful tool because it provides a different angle from which to view an active nucleus. □

Martin Ward is at the Institute of Astronomy, Cambridge University, Madingley Road, Cambridge CB3 0HA, UK.

Daedalus

A lesson from nature

MASONRY and concrete have excellent compressive strength but cannot reliably take tension. Any tensile load threatens catastrophic failure; the whole classical architectural tradition has evolved to keep every component of a building safely and permanently in compression. The reason, of course, is that masonry is full of cracks. Even a trivial tensile stress, amplified by the stress-concentration around a crack, may initiate brittle fracture.

So Daedalus hopes to emulate nature's way with concrete-like materials. Bones, teeth, tusks and the shells of molluscs are all made of highly brittle crystals of calcium salts. But they are bound together by a tough proteinaceous 'glue' whose slight yielding relaxes the stress-concentration around an incipient crack and makes it harmless.

Daedalus plans to exploit the way those tough and resilient siloxane polymers, the silicones, are made by the hydrolytic coupling of organosilane halides. He points out that these halides can bind chemically to a silica surface by a very similar reaction. So DREADCO's chemists are synthesizing various organosilane halides to spray onto, or to diffuse as vapour into, brick and concrete. They will bind on the siliceous surfaces of the cracks and fill them up with tenacious silicone resin. The masonry will be converted from a brittle porous solid to a tough and coherent one, as strong in tension as in compression. The treatment will probably only penetrate a few centimetres into the concrete. But since tensile failure begins at the surface of a component where the stress is greatest, silicone 'case-toughening' will still be effective.

Case-toughened concrete will revolutionize architecture. Steel and wood, so vulnerable to rust, fire and rot will no longer be needed: the tough new concrete will do it all. Its coherent silicone surface will even save it from the other curse of masonry — water-penetration and frost damage. Freed from the discipline of compression design, architects will doubtless rush to erect extravaganzas and monstrosities even more nightmarish or depressing than their current achievements. But on the positive side, many unique and graceful masterpieces of the past now crumbling into ruin or creeping into tensile instability could be rescued by *in situ* case-toughening (the leaning tower of Pisa comes to mind). And the new concrete may even invade other fields. Concrete chairs, tables and even grand pianos may become commonplace; waterproof and rustproof concrete ships may dominate the seas; and on a small but valuable human scale, repaying the debt to nature's methods, concrete false teeth.

David Jones

Acidification of Norwegian lakes

SIR—Brakke *et al.*¹ claim to have quantified acidification of Norwegian lakes by deposition of H_2SO_4 using a multiple linear regression equation relating strong acidity (as defined by Gran titration) to the independent variables TOC, SO_4^{2-} , and Ca^{2+} . Such quantification is not valid, however, because these three variables are not independent. Natural cycles of bases and acids (for example Ca^{2+} and TOC) in watersheds are perturbed by inputs of acid countering the effect of such input. Thus changes in acidity may be considerably less than predicted.

Henriksen, an author of ref. 1, has shown that concentration of Ca^{2+} and other base cations can significantly increase above background levels because of acidic deposition². Estimated increases are highly variable, but an increase equivalent to 40 per cent of estimated anthropogenic SO_4^{2-} input was put forward as an estimated regional mean value for southern Norway².

Organic acids (as represented by TOC) and SO_4^{2-} are not independent variables either. The same chemical principles dictating that inputs of H_2SO_4 increase dissolution of mineral bases dictate that inputs of H_2SO_4 decrease dissolution of terrestrial humic acids³. This pH-dependent solubility, along with other interactions⁴, results in replacement of humic acids by H_2SO_4 without the expected corresponding large change in acidity which is equivalent to estimated change in SO_4^{2-} .

Brakke *et al.*¹ cite palaeolimnological data for southern Norway and state that "these results suggest some reduction of TOC during acidification". Actually, these data indicate that pre-industrial levels of TOC were 2–3 times greater than present levels and loss of TOC was concomitantly accompanied by a surprisingly small change in pH⁵. A compilation of palaeolimnological data show that 9 of 10 acidic (pH < 5.5) lakes examined in southern Norway were always acidic⁶.

A similar compilation, which also includes water chemistry data, indicates that 12 of 14 acidic lakes in the northeastern US were acidic in pre-industrial times. Ten of these 12 lakes are now clear water and SO_4^{2-} -dominated and yet most show little or no measurable acidification.

These data, combined with reports of widespread loss of humic colour in the Adirondacks as well as in southern Scandinavia, plus theoretical considerations and experimental data, led the National Academy of Sciences report to suggest that organic acid buffering is important. Norton, an author of ref. 1, is co-author of this report⁷.

Reported values for conductivity in Tables 1 and 2 (ref. 1) are erroneously low. For example, conductivity for Lake G-21 in Table 2 is reported to be 2.36 μS (rather than $\mu\text{S cm}^{-1}$) when conductivity for H^+ alone is calculated to be 11.06 $\mu\text{S cm}^{-1}$.

Organic acids principally influence pH as weak acids not strong acids. Yet the authors assume that organic acidity is independent of pH (as is shown in the earlier

discussed multiple linear regression equation) and their running a linear regression of anion deficit on TOC: $\text{DIFF} = 4.06 \text{ TOC} - 0.96$.

Brakke *et al.*¹ incorrectly quote the results presented by myself and Frink³, "lakes were acidic simply because of processes occurring in acid soils". But we concluded, "acidification by acid rain is superimposed upon long-term acidification". An example was given where acid rain was believed to be the principal cause of acidification; however, we felt that "natural soil formation was often more important than acid rain in determining the acidity of lakes and streams".

EDWARD C. KRUG

Illinois State Water Survey,
2204 Griffith Drive,
Champaign,
Illinois 61820,
USA

Sequence-based phylogeny in eukaryotic genomes

SIR—Jared Diamond¹ compares the most parsimonious primate phylogeny deduced from DNA sequences of the η -globin gene with that deduced from data on hybridization of total single-copy DNA. The two methods yield the same branching pattern. "A moment's reflection shows why this conclusion was inevitable", he says, pointing to the fact that both methods actually measure the same thing, namely the number of mismatches in heteroduplex DNA, and that the two methods will converge on the same result as the size of the sequenced fragments increases.

His conclusion is obviously true when the length of the compared sequences approaches the total length of a genome. When studying the evolution of eukaryotic genomes, however, we are far from this ideal, and one must make a choice of sequences to compare. This choice, for two different reasons, could strongly influence the final result.

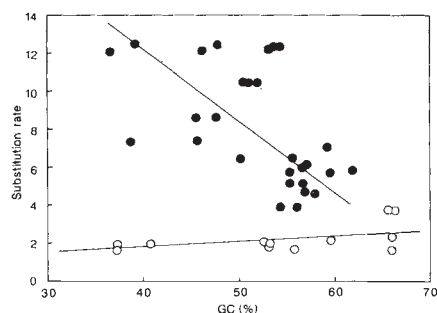
First, hybridization data reflect the number of mismatches in heteroduplexes prepared from bulk DNA, the majority of

which is non-coding. Mutations occurring during evolution in non-coding DNA have, in general, been neutral as far as their selective value is concerned. Therefore, they could have been indiscriminately accepted and fixed.

The situation is different in coding sequences of DNA (the major component of data banks). Selection would have prevented the fixation of some mutations, so that the apparent rate of accumulation of mutations is lower in the coding than in the non-coding part of the genome. Hybridization and sequencing data will only give convergent results if the ratio of these two rates of accumulation of mutations has been constant during evolution. But there are examples showing this not always to be the case².

A second problem, which is both more serious and only recently recognized, is caused by variation in the rate of accumulation of mutations³ with the chromosomal localization of the DNA sequence. In primates, GC-rich sequences located in Giemsa-negative chromosomal bands (see ref. 4) accumulate neutral mutations with approximately the same rate as the AT-rich sequences located in Giemsa-positive chromosomal bands. In rodents, however, while GC-rich sequences evolve at a rate similar to that of the equivalent human sequences, AT-rich sequences evolve four times as fast as their human equivalent (see figure). It is likely that the high rate of accumulation of mutations in AT-rich sequences in rodents as well as their compositional bias is caused by less efficient DNA repair in these regions of their genome⁵.

Whatever their origin, differences in the rate of accumulation of mutations cause sequence-based phylogeny to be dependent on the choice of sequence



Relationship between the rates³ [(silent substitutions/year $\times$ site) $\times 10^{-9}$] of the silent codon substitutions in several primate (open circles) and rodent (black circles) genes and their G+C content (from the EMBL sequence data bank).

1. Brakke, D.F., Henriksen, A. & Norton, S.A. *Nature* **329**, 432–434 (1987).
2. Henriksen, A. *Acid Rain Res. Rep.* no. 1 (NIVA, Oslo, 1982).
3. Krug, E.C. & Frink, C.R. *Science* **221**, 520–525 (1983).
4. Krug, E.C., Isaacson, P.J. & Frink, C.R. *J. Air Pollut. Control Ass.* **35**, 109–114 (1985).
5. Davis, R.B., Anderson, D.S. & Berge, F. *Nature* **316**, 436–438 (1985).
6. Scruton, D.A., Elner, J.K. & Howell, G.D. *Can. Tech. Rep. Fish. Aquat. Sci.* no. 1521 (Dept. Fish. Ocean, St John's, Newfoundland, 1987).
7. Charles, D.F. & Norton, S.A. in *Acid Deposition: Long Term Trends* 335–431, 482–506 (National Academy, Washington, DC, 1986).

analysed and so may not always produce the same phylogeny as that based on hybridization. The fact that the phylogenies shown by Diamond are alike, and so probably correctly represent the pattern of primate evolution, must mean that mutation accumulation rates in different chromosomal regions of primates have been similar.

J. FILIPSKI

*Institute Jacques Monod,
2 Place Jussieu — Tour 43,
75251 Paris,
France*

1. Diamond, J.M. *Nature* **332**, 685–686 (1988).
2. Beintema, J.J. & Campagne, R.N. *Molec. biol. Evol.* **4**, 10–18 (1987).
3. Li, W.H. & Tanimura, M. *Nature* **336**, 93–96 (1987).
4. Goldman, M.A. *et al. Science* **226**, 686–692 (1984).
5. Filipski, J. *FEBS Lett.* **217**, 184–186 (1987).

Scanning optical microscopy of low-contrast samples

SIR—A recent letter¹ by Rothenhäusler and Knoll considered the microscopic imaging of low-contrast samples and presented images of molecular films formed by the technique of surface-plasmon microscopy. It was stated that monolayer films do not provide sufficient contrast to be visible with either phase-contrast or Nomarski microscopy. We have examined the imaging of monomolecular films using various techniques of scanning optical microscopy² and our results lead us to dispute this claim. In particular the differential phase-contrast method³, using a detector split into two halves, is especially suitable for such applications. Here we show that the technique provides high sensitivity to specimens of low contrast.

The specimen used was a monolayer film of ω -trisenic acid ($\text{CH}_2=\text{CH}_2(\text{CH}_2)_{20}\text{COOH}$) deposited onto a glass substrate. The thickness of the film, 3.15 ± 0.05 nm (ref. 4), is a function of the length of the molecule and the angle at which the molecules are aligned. The

optical system used is shown in Fig. 1. Illumination is provided by a HeNe laser ($\lambda=632.8$ nm) and the objective lens is of numerical aperture 0.5. The system has both a transmitted-light split detector for differential phase contrast and a reflected-light detector which can be used either for bright-field imaging or, with a pinhole placed in front of it, for confocal imaging⁵.

The output of either detector can be connected to an LSI-11 computer to be sampled and quantized over a 256×256 grid by an eight-bit analogue-to-digital converter and stored in a frame memory. Images are generated by scanning the specimen mechanically in an xy raster through the focused spot. Imaging time is ~ 3 s per frame.

A differential phase contrast image of the edge of the film is shown in Fig. 2: the glass substrate is on the left and the monolayer on the right. The dark vertical line between these two areas is the edge of the film. To the right of this line can be seen areas where the film has not adhered to the substrate during dipping. These areas are clearly depressions: the sign of the gradient at their edges is bright on the left, indicating a downward slope, and dark on the right, indicating an upward slope. Not only does this method image the 3-nm phase steps easily, but it also shows up smaller detail on the surface of the slide and of the film.

To demonstrate the sensitivity of this technique to phase variations a confocal reflection image was observed for comparison. Despite a 25-fold enhancement in electronic contrast the confocal image showed the film edge only very weakly and also contained much less information about the surface structure of the film. A reflected light bright-field image exhibited even less detail than the confocal image despite a 30-fold contrast enhancement.

In Fig. 2 the dark line representing the edge of the film has a width of about $2 \mu\text{m}$. If the edge is assumed to be a linear

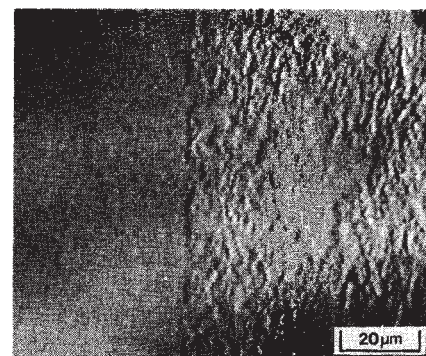


Fig. 2 A differential phase contrast image of a single monolayer.

ramp rising 3 nm vertically in $2 \mu\text{m}$ horizontally, the normalized intensity signal is estimated² as 0.002, agreeing well with the observed value. The theoretical contrast in the edge in a bright-field image is also 0.002, and in a confocal image it has twice this value². These also agreed well with observed values, demonstrating the improved sensitivity of confocal microscopy to small phase changes.

The differential phase contrast method has been demonstrated to detect phase gradients of about 1 in 1,000, which is equivalent to a height change, over a resolution element, of less than 1 nm. One reason for this high sensitivity is that mechanical scanning avoids the appearance of mottle in the image⁵. The sensitivity can be improved further by using an annular split detector⁶, giving an increase in sensitivity of about a factor of 10, or by using two photomultipliers or photodiodes instead of the solar cell detectors actually used. If small photodiodes are used it is then necessary to focus each half-beam individually onto its corresponding photodiode.

C.J.R. SHEPPARD
D.K. HAMILTON
H.J. MATTHEWS

*Department of Engineering Science,
Oxford University,
Parks Road,
Oxford OX1 3PJ, UK*

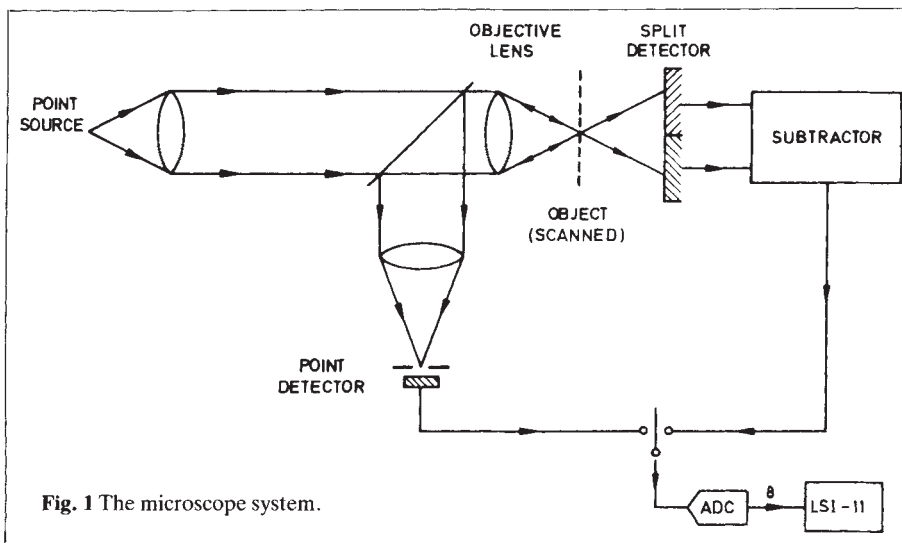


Fig. 1 The microscope system.

1. Rothenhäusler, B. & Knoll, W. *Nature* **332**, 615–617 (1988).
2. Sheppard, C.J.R. in *Advances in Optical and Electron Microscopes* Vol. 10 (eds Barer, R. & Cosslett, V.) 1–98 (Academic, London 1987).
3. Hamilton, D.K. & Sheppard, C.J.R. *J. Microsc.* **133**, 27–39 (1984).
4. Bonnerot, A., Chollet, P.A., Frisby, H. & Mollet, M. *Chem. Phys.* **97**, 365 (1985).
5. Allen, R.D. & Allen, N.S. *J. Microsc.* **129**, 3–17 (1983).
6. Hamilton, D.K., Sheppard, C.J.R. & Wilson, T.J. *Microsc.* **153**, 275–286 (1984).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

From eugenics to genocide

Benno Müller-Hill

Racial Hygiene: Medicine under the Nazis. By Robert N. Proctor. *Harvard University Press: 1988. Pp. 414. \$34.50, £27.95.*

Race Hygiene and National Efficiency: The Eugenics of Wilhelm Schallmayer. By Sheila Faith Weiss. *University of California Press: 1988. Pp. 245. \$35, £20.50.*

WHEN scientists think about the history of science, they like to think about the lives of a few great men and women. The greatness of such people is defined by the influence of their discoveries on present thinking. Fields which have ceased to develop, or which may even have disintegrated, are regarded as irrelevant — who in his right mind would care for their history? This view could be accepted if science were small and had just the one function of increasing knowledge. But science, and particularly biomedical science, has become a vast enterprise that has taken on the additional roles of social control and justification of political measures.

Robert N. Proctor has chosen to analyse these roles in his book *Racial Hygiene: Medicine under the Nazis*. How did medicine fare under the Nazis? Did it, one might ask, fare as badly as mathematics for example? Proctor's conclusion is that this was not so. Many of the intellectual and social foundations of the Nazi programmes were outlined by biomedical scientists long before Hitler came to power in 1933, and these scientists played a leading part in the initiation, administration and execution of Nazi racial policy.

So Proctor begins with a discussion of eugenics, or 'racial hygiene' as it was called in Germany. The founder of eugenics in Germany, Wilhelm Schallmayer, was neither antisemitic nor an advocate of the supremacy of the nordic race. But many if not most of the others who later joined the movement became, in time, antisemites and firm believers in nordic superiority. Proctor mentions that more than a dozen journals devoted to eugenics or its pro-nordic, antisemitic variant, racial hygiene, were published in Germany before 1933.

The relative smoothness with which the biomedical establishment coped with the transition in 1933 is documented in an appendix: the editorial boards of most medical journals (Proctor lists 147) remained virtually unchanged between 1932 and 1934! The majority of the German

medical profession gratefully rejoiced when their Jewish colleagues were kicked out. They got their patients. In 1933 the average taxable income of a German physician was 9,280 RM; by 1938 this had grown to 14,900 RM. Proctor also quotes Michael Kater's data and adds new material to show that about 45 per cent of the German medical doctors ultimately joined the Nazi party — a step which was

350,000 to 400,000 persons, was at first welcomed internationally and cites the opinion of an anonymous member of the Eugenics Records Office in Cold Spring Harbor that Hitler "should be made an honorary member of the Eugenics Records Office" — which never happened. He documents the enthusiasm with which German gynaecologists explored new methods of sterilization, and recalls the favourable reaction of the German medical press to the Nuremberg Laws: "German academic Journals also made much of the fact that the American black-white miscegenation laws were generally much stricter than Germany's own Nuremberg Laws". He also discusses medical quarantine, which was used as a measure to speed up the Holocaust in the ghettos in Poland.

Further chapters cover what Proctor calls the "control of women", "the organic vision" and "medical resistance". It is this very diversity which makes parts of the book tough going. Proctor obviously had to rely on secondary sources to deal with the mass of different problems, and some of these sources are better than the short versions he gives. Furthermore it is difficult and perhaps impossible for the average reader to take in all these various issues — for example how the medical establishment dealt with midwives, and quacks, its fight against tobacco, DDT and alcohol and for good bread — and then read about genocide. True, reality is diverse, but Proctor's book may make some readers wish to go back for good to the simple history of the heroes of science and medicine.

I found most of the text free of misprints and factually correct, but there is one major mistake. The records of the Wannsee Conference were not found in the files of a member of the SS but in those of officials of the German Foreign Office. They — and only they — had forgotten to destroy this most damaging evidence. The book is well produced. The well chosen illustrations are revealing and the tables of medical journals and chairs of racial hygiene are illuminating.

In her book, *Race Hygiene and National Efficiency*, Sheila Weiss traces the history of the recent genocide in Germany back to its scientific origins. Was eugenics in Germany fundamentally different from Anglo-Saxon eugenics? This myth, which was propagated in the United States and Britain, is here destroyed. ▶

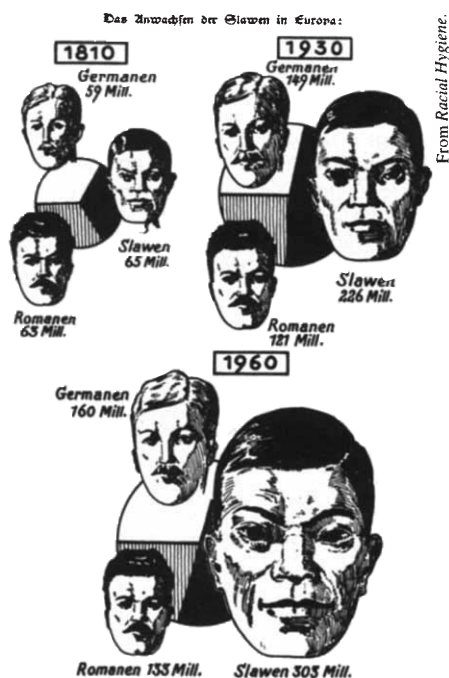


Prize-winners of the "best Nordic head contest" from Volk und Rasse, 1927. (No first prize was awarded for the women's contest.)

certainly not necessary for a normal career. Proctor concludes: "The general point to be made is that under financial pressures, solutions that seem extreme or inhumane [in one set of circumstances] may become banal to a large number of people in another [set of circumstances]". This important point seems to have been overlooked by many observers — just possibly because of its utter banality.

Proctor devotes extensive sections to the sterilization, the killing of the insane and the antisemitic measures. He points out that the sterilization law of 1933, which led to the sterilization of some

From *Racial Hygiene*.



From Racial Hygiene.

"Fertility and race, the growth of the Slav in Europe", from O. Helmut's *Volk in Gefahr*, 1934.

Weiss demonstrates convincingly that Wilhelm Schallmayer was opposed to the ideas of nordic supremacy and antisemitism. She calls Schallmayer's position "nonracist" but points out that as far as blacks were concerned Schallmayer was, like almost everybody else in the field, an outright racist. She recounts the story, often told in German publications, of how he won the competition set up in 1900 by Krupp — "What can we learn from the theory of evolution about internal political development and state legislation?" — and outlines the goal of the competition. Its object was to wrest darwinism out of the hands of the labour movement and use it for eugenic legislation to increase the efficiency of the German capitalist state, just as the Krupps had made their steel mills more efficient. Weiss shows that, internationally, the aim of all the various brands of eugenics was to increase the efficiency of the state. But the later fatal development which, through the efforts of Baur, Fischer, Lenz and Rüdin, occurred in Germany in the 1920s, has no place in her story because Schallmayer, her hero, died in 1919. I note a curious omission from this book: Nietzsche, who probably had a fundamental influence on Schallmayer's thinking, is not even mentioned.

Up to a point, these two volumes complement each other. But the book which describes the history of human genetics and its international darwinistic implications has yet to be written. □

Benno Müller-Hill is a Professor in the Department of Genetics, University of Cologne, D 5000 Köln 41, FRG. An English translation of his own book, *Murderous Science: Elimination by Scientific Selection of Jews, Gypsies and Others, Germany 1933-1945*, has just been published by Oxford University Press.

Information on propagation

Barry J. Everitt and Martin H. Johnson

The Physiology of Reproduction, Vols 1 & 2. Editors-in-chief Ernst Knobil and Jimmy D. Neill. Raven: 1988. Vol. 1 pp. 1,390; Vol. 2 pp. 1,021. \$362.50, £214.50

With these two volumes, Knobil and Neill hope to "fill a need for a comprehensive, scholarly treatise on the physiology of mammalian reproduction". The scope of the work is as broad as would be expected from the title. Volume 1 is divided into three sections ("The Gametes, Fertilization and Early Embryogenesis"; "The Reproductive Systems", itself split into subsections devoted to the female and the male; and "The Pituitary and Hypothalamus"), Vol. 2 into two ("Reproductive Behavior and its Control"; and "Reproductive Processes and Their Control", the latter addressing topics such as puberty, cyclicity, pregnancy and reproductive senescence). The books are not profusely illustrated; of the illustrations that are included, many of the photographs are poor and the line illustrations rather thick textured.

The volumes are certainly scholarly. Most of the authors are established researchers who have made original contributions to the areas they review. Glancing at the table of contents gives the impression that the work is also comprehensive; the chapter titles bear little evidence of obvious omissions — although fetal physiology is perhaps one.

On reading further, some shortcomings become apparent. Although the editors assert that "it has been left to the reader to ascertain the similarities and differences" in reproductive processes among mammalian species, they have not always taken the trouble to ensure that the necessary, comparative information has been included. This is especially the case so far as human reproduction is concerned. Where a chapter covers the range of species, the human male or female are often mentioned only in passing or not at all. Where a group of chapters covers a single topic from a species perspective, for example "Puberty", there are discussions of the rat, the sheep, primates — but not the human being, despite the existence of a considerable clinical experimental literature on the subject.

The problem can be illustrated by reference to one contribution in particular, that by Pfaff on female reproductive behaviour. It is excellently written and reviews the author's own work on defining the neuroendocrine mechanisms which underlie the lordosis reflex in female rats. This important research, conducted by

Pfaff over many years, has enabled the dissection, with great precision, of the somatosensory pathways involved in the reflex, the ways in which ovarian hormones facilitate it, the neurochemical mechanisms involved and even the intracellular, molecular events set in train by oestrogen's action on nerve cells. But readers turning to this chapter to find out about the neuroendocrine control of sexual behaviour will be disappointed indeed, unless they are especially interested in the section on cellular principles. The hormonal basis of sexual behaviour in female primates, including women, is very different to that in the rat, and lordosis is a sexual reflex not even encountered in these species. This sort of failing is particularly irritating given the unnecessary duplication of material elsewhere. For example, the oviduct and its functions are covered in two or three places, as is epididymal physiology.

However, many of the contributions are superb. In addition to that by Pfaff, those by Sachs, Numan, Plant, Turek, Challis, Pedersen and Setchell, among others, are informative, contemporary accounts of important topics. So although this is not a comprehensive textbook of reproductive physiology, it is an excellent collection of reviews which covers, albeit in a variable and idiosyncratic way, a wide range of areas of this broad field of research. True, many chapters are pitched at a high level, and to appreciate much of the information in the books it is necessary to have a considerable background knowledge not just of the subject as a whole, but of the specific topic of each chapter. For example, the plunge into the complexities of sex determination on the first page of Chapter 1 with a passing, unexplained reference to "Xsr" is going to lose many inexperienced readers rapidly.

Postgraduate students will find these volumes demanding. Experienced, post-doctoral researchers will find them, as we did, extremely valuable and informative. This work will be essential for every laboratory engaged in reproduction research and every library in scientific and medical environments where such research is being conducted. It is an impressive enterprise. □

Barry J. Everitt and Martin H. Johnson are in the Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK.

New editions

- *Plant Molecular Biology*, 2nd edn, by D. Grierson and S. N. Covey. Publisher is Blackie/Chapman and Hall (New York), price is hbk £27, \$45.75; pbk £12.95, \$21.95. For review see *Nature* 314, 41 (1985).
- *Microcomputers and Laboratory Instrumentation*, 2nd edn, by D. J. Malcolm-Lawes. Publisher is Plenum, price is £27.95, \$47.40. For review see *Nature* 311, 313 (1984).

Rock of ages

Euan G. Nisbet

Science and Earth History: The Evolution/Creation Controversy. By Arthur N. Strahler. *Prometheus, Buffalo, New York: 1988. Pp.552. \$39.95. In Britain distributed from 10 Crescent View, Loughton, Essex IG10 4PZ, £29.95.*

SOME of the strongest words in the Bible are reserved for those who mislead humanity by speaking falsehood and claiming that it is truth. These false shepherds — St Jude calls them clouds without rain — are those who by perverting the truth manage to lead the faithful astray. One of the most successful of the stumbling blocks, or obstacles to truth, is the pseudoscience of 'creationism', which is exploited by the legion of false prophets inhabiting the television screens of North America. The claim of 'creation scientists' is that geologists have misunderstood the Earth, and that the geological record is fully compatible with a strictly literal reading of the Bible.

Europeans, who for the most part live in a non-Christian society, often express contempt for North American culture that needs to concern itself with this claim of creation science. Why not simply ignore the charlatans and the misguided peasants who follow them? Such a view takes no account of the real strength of creationism and its record of profound influence on the political process. It also forgets the much sadder European history of the opposite phenomenon, of how the perversion of pseudo-darwinism provided the intellectual underpinning of National Socialism and the Nazi onslaught on Jews, gypsies and mental 'defectives.'

There are two defences against creationism. One is theological, the other scientific. In this long and complex book, A. N. Strahler has chosen the latter course. Creation science is an extraordinary mélange of bent evidence, misinterpretation and subtle selection of data that is based not on truth but on a value which our society places higher: credibility. It is a pseudoscience steeped in the ethics of the law courts or the opinion polls — that argument wins which is most 'credible' (that is, most loudly shouted). Strahler's book is itself not only a shout against creationism but, more important, a detailed point by point rebuttal of the farrago of untruth.

The book's length and complexity reflect the diversity of the arguments of creation science. First there is the point that in order to recognize pseudoscience we must be able to identify real science — which is nothing less than the search for truth about the natural world — and Strahler gives a long account of the contrasts

between science and pseudoscience, and of the basic tenets of creation science. This section is followed by detailed expositions of the contrasts between the way in which a 'mainstream' scientist views the history of the Earth and Solar System and the way in which the same evidence is seen by creationists. This naturally leads on to the age of the Earth, to Noah, Adam and Eve, and the interpretation of the fossil record.

Strahler's exposition is careful, thorough and clearly written. If at times the book appears to be disordered, that is a result of the disorderly fabric of creationism. The author pulls this fabric apart, thread by thread, and with better yarn weaves an excellent account of the fabric of mainstream natural science, especially geology. Of course, 'true' science does not stand still — we find deeper truth, or we discover flaws in our fabric — and there are sections of the text which have the dated flavour of an old textbook. But, in general, Strahler provides a fine exposition of the history and properties of the natural world. For those who, like myself, teach introductory classes in science, *Science and Earth History* will be a splendid source of material in the defence of truth, and a great help to those of us who still hold to the Jewish or Christian faith.

There is, however, the other defence against creationism which is less often used, though it is immensely effective. This is that creationism is, in Judaeo-Christian terms, nothing less than heresy — not only because it denies truth, but because it is in opposition to the long line of Jewish and then Christian theology that began in the Jewish community in Alexandria and was developed in more detail by St Augustine in his magnificent commentary on Genesis. Karl Barth, in his comment on Romans 8.28, puts the theological position best: "we lend our support

to all honest, secular, scientific and historical research; but we dissociate ourselves from every semi-theological interpretation of Nature and History".

By misinterpreting the English of the King James Version, creationism attacks the true meaning of both the Old and New Testaments. Augustine's discussion of time is incisive and strangely parallels Einstein's. Time, in orthodox Christianity, is not simply some consequence of the movement of a heavenly body. It is a created dimension. The 'days' of creation, like the Day discussed by the writer of the epistle to the Hebrews (Hebrews 3.13 and 4.7), are not measured by the rotation of the Earth. Augustine realized that the vision of Genesis lies outside time, in eternity. The creationists seek to trivialize God, to reduce the deity to the bounds of time: Augustine pointed out that the narrative of Genesis is a vision of truth that is concerned with something much higher than pseudoscience.

The orthodox Christian belief is that the Sun has not yet set on the seventh day of creation, despite what the creationists say, and there is a fine exposition of conservative belief in Derek Kidner's commentary on Genesis (Tyndale Press, 1973). For those whose university classes are plagued with creationism, Strahler's book, together with Kidner's essay and St Augustine's commentary on Genesis, may serve to protect the truth. Students with problems should be advised to read these, and of the television evangelists to watch only Billy Graham. *Science and Earth History* is a valuable rejection of the creationist heresy, for which many of us will be deeply thankful. □

Euan G. Nisbet is a Professor in the Department of Geological Sciences, University of Saskatchewan, Saskatoon, Saskatchewan S7N 0W0, Canada.



Worshipping material things — sacrifice is offered to eight goddesses of silkworms. The central importance of silk production in China was expressed both in official sericultural ceremonies and the popular sacrificial cult of the silkworm breeders. The illustration, from the Nung Shu (Book on Agriculture) of 1313 is taken from *Science and Civilization in China*, Vol V:9, Textile technology, by Dieter Kuhn. Publisher is Cambridge University Press, price is £60, \$110.

The importance of relaxation

Myron Lecar

Dynamical Evolution of Globular Clusters. By Lyman Spitzer, Jr. *Princeton University Press*: 1988. Pp. 180. Hbk \$35, £20.50; pbk \$14.50, £8.50.

OUR Galaxy contains close to 200 globular clusters. A variety of age indicators (chemical composition, orbits and the age of the oldest stars still on the main sequence) confirm that they were born early in the life of the Galaxy, about 10^{10} years ago. The study of the dynamical evolution of these clusters is irresistible to statistical physicists because this age is much longer than their 'relaxation time' (that is, the time for the orbits of individual stars to be appreciably perturbed by two-body encounters). Most of the well-observed globular clusters have relaxation times of between 10^8 and 10^{10} years. By comparison, the relaxation time of the Galaxy exceeds 10^{14} years. The present structure of most globular clusters should show few traces of their initial structure.

In much the same way as stars, globular clusters evolve through a sequence of quasi-steady states, which are steady state solutions of the collisionless Boltzmann equation. There is no equilibrium state; the system always increases its entropy by forming a tightly bound core (to increase the negative binding energy) while expanding a halo or unbinding stars (to occupy a larger volume in phase space). But, because the time for a star to orbit the cluster is much shorter than the relaxation time, the structure is almost stationary on a dynamical time scale.

For most of their lives, the evolution of globular clusters is well traced by solutions of the Fokker-Planck equation. Professor Spitzer has been a leader in the simulation of solutions to the Fokker-Planck equation by Monte-Carlo techniques. In this book, he discusses his own methods and results, and gives fair attention to the other major practitioners of the Monte-Carlo technique (Henon, Shapiro and Teukolsky) and to the direct solution of the Fokker-Planck equation by Cohn.

Spitzer's skilful illumination of the physical phenomena that drive evolution means that this book will become a classic. Using idealized models, he illustrates the phenomena of evaporation of stars from an isolated cluster; from a tidally limited cluster; mass stratification; and the collapse of the isothermal core because of the gravo-thermal instability. Heating of the halo by compressive shocks (as the cluster passes through the disk of the galaxy) and tidal shocks are also treated.

The book concludes with a discussion of the current puzzle in globular cluster evolution. It is now firmly believed that the evolution proceeds to a thermal runaway in the core. After a finite time, the density of an innermost region, well within the isothermal core, tends to infinity. Furthermore, many globular clusters should have undergone such core collapse some time ago. Theorists are now asking: "What happens after core collapse?" A plausible suggestion is that when the density becomes very high, binary stars are formed which, through close encounters with a third star, increase the kinetic energy of the core, at the expense of increasing the (negative) binding energy of the binary. This process is another example, on a smaller scale, of

the same instability that prompted the core collapse in the first place. The influx of kinetic energy into the core brakes the collapse of the core and causes it to expand. Some simulations indicate a cycle of collapse followed by re-expansion.

These phenomena are currently being studied. For an astrophysicist who wants to join the endeavour, Spitzer's book will be indispensable. Here is to be found a complex theoretical structure assembled by a master craftsman. The writing style is compact, which probably makes the book unsuitable for an introductory course in stellar dynamics, but this is an excellent text for a graduate course or seminar. □

Myron Lecar is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

Getting results

F.W. Bullock

How Experiments End. By Peter Galison. *University of Chicago Press*: 1988. Pp. 330. Hbk £31.95, \$47.95; pbk £11.95, \$17.95.

A DISCUSSION of Einstein's work as an experimental physicist is hard to come by. For instance, Ronald W. Clark in his biography makes no mention of the Einstein/de Haas experiments which were aimed at determining the g-factor of the electron. That these took place at about the same time as Einstein was struggling with the General Theory of Relativity, and was anxiously (or confidently perhaps) awaiting the outcome of the ill-fated Freundslich solar eclipse expedition to the Crimea, would appear to testify to the range of Einstein's abilities and interests.

It is more likely, however, that Einstein was not an active participant at the laboratory bench in the experiments which form part of the first case study in Galison's exhaustive search for the factors determining at what point an experimental study can be said to have achieved its objectives and can therefore be concluded. The other two series of experiments analysed in impressive detail are those (mainly cosmic-ray) experiments which eventually led to the discovery of the muon, and the accelerator experiments which resulted in the discovery of the neutral current process in weak interactions. In each of these studies the author presents the many levels of theoretical supposition that determined the experimenters' activities in deciding both the objectives of their experiments and the way the data were analysed and eventually presented.

The depth of Galison's research cannot be denied. In the case of the discovery of the weak neutral current he has been able to interview many of the researchers

involved. Although his account would not always agree in fine detail with every memory of those who participated in the experiments, there is no doubt that he has understood very well the factors which determined events, from the initial speculation that an effect might be present to the final publication reporting the existence of neutral currents.

Some 80 per cent of the book is taken up by a comprehensive and scholarly account of theories, apparatus, experimental method and, in some cases, the personalities and prejudices of the people involved. Although providing ample evidence of the author's industry, and some interesting reading, it is often difficult to perceive the relevance of all the minutiae included. Indeed, on reaching the final chapters the reader may well be anticipating that his reward for persevering will be the appearance of a philosophical statement of some depth. In this respect, the conclusion is anti-climactic. Much of what has appeared before is repeated and the attempt to parse the events into a braudel-type structure of long-term, medium-term and short-term constraints does not illuminate in any novel way the question of how experiments end.

In spite of this, any scientist interested in the philosophy or sociology of the experimental method will find the book worth reading. It demonstrates not only the many achievements of scientific research but also the many failings of those who pursue it. At the end, however, one must still wonder not so much how or why experiments end, but why the author chose these particular experiments to analyse. He gives us no reason to believe that they are typical of experimental physics and certainly leaves open the question of whether they are in any way representative of science in general. □

F.W. Bullock is a Reader in the Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK.

Junction effect interactions in methanol synthesis catalysts

J. C. Frost

BP Research Centre, Sunbury-on-Thames, Middlesex TW16 7LN, UK

The mechanism of catalysis of methanol synthesis is controversial. Here it is proposed that the minute Schottky junctions at the interface between metals and oxides in the catalysts affect the surface chemistry of the oxides in a way that correlates with catalytic behaviour. This theory can predict the existence and properties of new catalysts.

METHANOL synthesis is a process of major importance for the chemicals industry. It is currently achieved by passing syngas (a mixture of carbon monoxide, hydrogen and carbon dioxide) over a copper/zinc oxide catalyst developed in the 1960s. Other oxides and metal/oxide combinations are also effective, but in all cases the productivity (production of methanol per unit volume of catalyst per unit time) of the individual metals or oxides is some orders of magnitude lower than the metal/oxide combinations. The explanation for this remarkable synergy has remained controversial.

There has been much debate¹⁻¹³ about the possible presence and potential importance of small amounts of non-metallic copper in the copper/zinc oxide catalyst, but there is now common agreement that the active catalysts consist almost entirely of copper metal and zinc oxide. Suggestions that there is extensive dissolution of the copper in the zinc oxide⁴⁻⁶ are not supported by recent *in situ* EXAFS (extended X-ray absorption fine structure) and X-ray photoelectron spectroscopy experiments⁸⁻¹⁰. Very recently the existence of any special synergy between the copper and the zinc oxide has been denied. In the presence of CO₂ some of the available metallic copper surface appears to be oxidized, and a large body of experimental work¹⁻³ has been interpreted as proving that this oxidized metal surface is the active catalytic component.

It is proposed here that the synergy is real and that its origin lies in the perturbation of the oxide defect equilibria by the metal/oxide junctions in the catalyst to give a dramatic (10^3) rise in the surface concentration of ionized oxygen vacancies. The calculated vacancy concentration is shown to be directly related to the catalytic productivity. This theory of junction effects (which owes much to some early theories of catalysis¹⁴) predicts the response of a wide range of methanol synthesis catalysts (including the copper/zinc oxide catalyst) to changes in feed gas composition.

Catalyst structure

Measurements made at the BP Research Centre show that copper/thoria catalysts exist solely as copper metal and thoria under reaction conditions. The copper is in two forms: large (>100 Å diameter) and very small (7-15 Å diameter) particles. The small copper particles, representing up to 95% of the total copper, are undetectable above the background in X-ray diffraction traces (M. Soar and A. G. Mitchell, unpublished observations), but are clearly seen in EXAFS experiments (P. Meehan, unpublished observations). The EXAFS analysis of samples reduced *in situ* and reacted with syngas shows no white line from oxidized copper, and the line shape can be fitted to metallic copper with reduced coordination spheres. The mean copper particle size is ~12 Å. Electron microscopy, including scanning transmission, measurements (D. Imeson, unpublished observations) are best interpreted as a bimodal distribution of copper particles encapsulated by thoria. The energy-loss spectroscopy copper absorption edge again reveals no sharp white-line resonances from oxidized copper anywhere in the sample. Low-temperature

adsorption followed by temperature-programmed desorption of CO, together with N₂O chemisorption experiments (A. G. Mitchell and P. Aukett, unpublished observations) reveal very little free copper metal surface (1-3 m² g⁻¹ catalyst) in comparison with the 200-300 m² g⁻¹ expected from such highly dispersed copper. The possibility that the chemisorption and electronic properties of these small metal particles differ from those of the bulk metal is not supported by measurements on small metal particles produced by laser vaporization¹⁵. The simplest interpretation of these experiments (the details of which will be reported elsewhere) is that the oxide encapsulates the small copper particles to give a very high metal-to-oxide interfacial contact area in the catalysts.

The situation in the copper/zinc oxide catalysts is controversial, but examination of literature data shows that a similar bimodal copper distribution exists with the small copper particles, accounting for perhaps 45% of the copper^{5,10}.

The interaction described below shows how metal particles, even when encapsulated by the oxide, can cause important changes in the chemistry at the oxide surface.

Junction effects

Methanol synthesis conditions are highly reducing and the metal components are stable as bulk metals, whereas the oxides will contain doubly ionized oxygen vacancies ($[V_O^{++}]$) as the dominant simple defect^{16,17}. The oxygen partial pressure ($p(O_2)$) under methanol synthesis conditions is determined by the CO₂/CO gas ratio because of the equilibrium $CO_2 = CO + \frac{1}{2}O_2$. In the absence of the metal, the important defect equilibria are oxygen vacancy creation and subsequent ionization. Using the accepted constants^{16,17} for zinc oxide, the calculated concentration of $[V_O^{++}]$ is in the region of 5×10^{16} cm⁻³ for ZnO under a 0.1 CO₂/CO gas ratio at 523 K. If the surface is defined arbitrarily as the 10 Å layer adjacent to the gas-solid interface, then the surface concentration is perhaps 10^9 - 10^{10} cm⁻². Computational studies (P. Tasker and J. Harding, unpublished observations) show that the energy of oxygen vacancies at the thoria surface is similar to that of vacancies in the bulk, suggesting that strong vacancy segregation does not occur.

If a metal such as copper is placed in contact with the oxide (as shown in Fig. 1a for copper/thoria) another equilibrium is established in which the electrons produced by ionizing the oxygen vacancies are distributed between the oxide conduction band states and the states available at the metal Fermi level (E_F). The conduction band edge (E_C) for these oxides is significantly higher in energy than the metal Fermi level, so there is a net transfer of charge from the oxide to the metal.

One of the main contributions to the enthalpy of formation of an oxygen vacancy is the energy required to raise two electrons to the oxide conduction band. Allowing these electrons to lower their energy by moving to the metal Fermi level effectively reduces the enthalpy of formation (H_v) of the doubly ionized oxygen vacancy by $2(E_C - E_F) = 2\phi(b)$, where $\phi(b)$ is the familiar Schottky barrier height used in semiconductor device

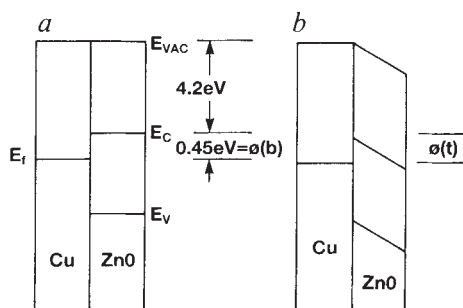


Fig. 1 Band diagrams for Cu/ZnO. (a) Unrelaxed Cu/ZnO contact. (b) Cu/ZnO contact after vacancy creation and electron transfer up to the band-bending limit. E_{VAC} , vacuum level; E_V , valence band edge; E_F , Fermi level; E_C , conduction band edge; $\phi(b)$, Schottky barrier height; $\phi(t)$, band-bending height.

physics¹⁸. This stabilization perturbs the defect equilibria in favour of increased oxygen vacancy concentrations.

There are constraints on the charge transfer process. At the low temperatures of methanol synthesis (~ 500 K) we expect oxygen to be removed only from the surface. The electrons must be able to move to the metal at a rate which at least enables the initial surface reduction process to occur. Calculations show that quantum mechanical tunnelling allows very rapid equilibration (10^{-6} s) if the oxide surface is about 30 Å from the metal oxide junction (J. W. Millar, unpublished). Many solid-state devices rely on tunnelling through such thin oxide layers. The rapid equilibration suggests that electron transfer rates between the oxide surface and the metal need not be rate-limiting for the catalytic process or for the initial formation of oxygen vacancies.

A second insuperable constraint is that charge transfer causes the oxide bands to bend downwards. Eventually the conduction band edge falls below the metal Fermi level and the electron transfer process stops (Fig. 1b). The band bending ($\phi(t)$) is proportional to the charge transferred [V_0^{++}] and the charge separation, and inversely proportional to the dielectric constant of the oxide. For oxide layers 10–20 Å thick, the band-bending limit for [V_0^{++}] is $\sim 1\%$ of the surface oxygen atoms.

The metal/oxide combination is overall charge neutral, and combining this condition with the defect equilibria described above allows a mass action equation for [V_0^{++}] to be derived:

$$[V_0^{++}] = (A/4)^{1/3} (M)^{2/3} p^{-1/6} (O_2) e^{(-Hv - 2\phi(b) + 2\phi(t))/3kT} \quad (1)$$

The value of $\phi(b)$ for copper/zinc oxide is 0.45 eV (ref. 19), A is in the region of 10^{29} – 10^{30} cm⁻³ atm^{1/2} (refs 16 and 17) and Hv is ~ 4.3 eV (ref. 17). M , the density of metal states available to receive electrons, is about equal to the number of metal atoms in direct contact with the oxide because the metal conduction electrons efficiently screen the rest of the metal particle from the metal oxide junction. M determines the absolute productivity of the catalyst and clearly depends on the catalyst microstructure. It will typically be $\sim 10^{21}$ (cm⁻³ oxide) for the catalysts described here, where the small metal particles are encapsulated by 5–50 Å oxide overlayers. Using these values in equation (1) gives concentrations of [V_0^{++}] between 10^{19} and 10^{20} cm⁻³ under a 0.1 CO₂/CO ratio feed at 523 K. This implies a surface concentration of 10^{12} – 10^{13} cm⁻² if the vacancies are concentrated at the surface of the thin oxide overlayers. The substantial increase over the concentration in pure zinc oxide is in accord with the increased productivity of the binary catalyst, and suggests a direct relationship between the perturbation of the defect chemistry and the catalytic behaviour.

This treatment of the defect equilibria is too simple, and complications such as different vacancy ionization states, vacancy interactions, and the association of the vacancies with the reactants will alter the detailed form of equation (1). Parameters such as M , A and so forth are also difficult to estimate

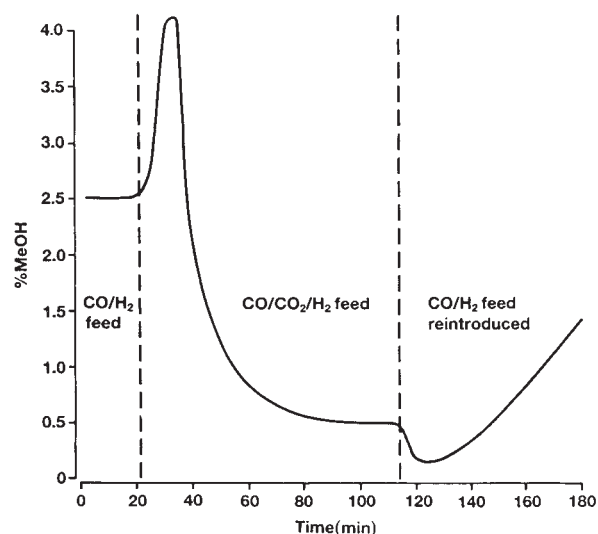


Fig. 2 Response of a Cu/ThO₂ catalyst's activity to changes in the feed gas composition.

with confidence. Nevertheless, the calculations are robust because the maximum concentration of ionized defects is determined by the band-bending constraint, rather than by the choice of parameters. The band-bending constraint dominates when $Hv - 2\phi(b)$ is below some critical value. The vacancy concentration is independent of oxygen partial pressure in this regime. For values of $Hv - 2\phi(b)$ above the critical value [V_0^{++}] becomes directly proportional to $p^{-1/6}(O_2)$ (that is $(CO_2/CO)^{1/3}$). Experimental values for the parameters for other metal/oxide combinations are sparse. But the Schottky barrier height for ionic semiconductor junctions is given by the difference between the metal workfunction and the oxide electron affinity¹⁸. Hence for a given oxide, $Hv - 2\phi(b)$ is determined by the workfunction of the metal. Similarly, $Hv - 2\phi(b)$ for a given metal depends on the reducibility and electron affinity of the oxide. The electron affinity of n -type oxides is close to the workfunction, whereas Hv can be estimated empirically¹⁷. Using these relations we expect $Hv - 2\phi(b)$ to increase with increasing free energy of formation (per oxygen atom) of the oxide.

Tests of the theory

We have made active methanol synthesis catalysts from silver, copper, palladium, platinum and gold metals with zinc, zirconia and thoria oxides (Table 1) in $\sim 1:1$ metal:oxide molar ratios, as well as the copper/zinc oxide catalyst. These catalysts have been prepared using a rising pH coprecipitation method with

Table 1 Comparative stable catalyst productivities

Catalyst	Productivity (kg h ⁻¹ m ⁻³)
Pd Zr	9.5
Cu Zr	24.1
Cu Zn	10.3
Pt Th	5.7
Pd Th	43.5
Cu Th	45.5
Au Th	10.0
Ag Th	10.3
— Th	<0.05
— Zr	<0.05

Comparative stable catalyst productivities at 10 bar, H₂/CO/CO₂ 2:1:0.01, 453 K, gas hourly space velocity 3,600 h⁻¹ after at least 24 h on stream. The productivity of the commercial catalyst under optimum feed ratios under these conditions is about 40–50 kg h⁻¹ m⁻³.

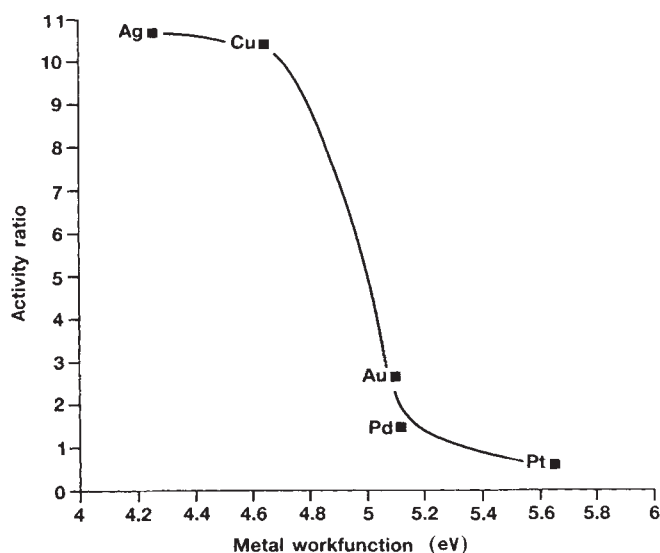


Fig. 3 Metal workfunction versus activity ratio, defined as initial activity/activity after exposure to CO_2 .

ammonium carbonate as the precipitant. The precursor salts were nitrates where these are stable and soluble. The gold catalyst was prepared from a chloride precursor. The precipitates were carefully washed with water, vacuum dried and calcined in nitrogen at temperatures between 723 K and 773 K. Copper, palladium, zirconia and thoria are recognized as useful components of methanol synthesis catalysts, but silver and gold are usually regarded as inactive. A copper-ceria combination is reported to be particularly active when prepared from the copper-cerium intermetallic compound²⁰. In this theory of methanol synthesis, the function of the metal is to promote the productivity of the oxide and, provided the metal does not perform unwanted side reactions such as the dissociation of carbon monoxide, its surface chemistry is of little importance. This is the origin of our success in promoting the base productivity of the oxides with such apparently unpromising metals as silver and gold. Provided the metal workfunction is large enough to form a Schottky junction with the oxide it will enhance the equilibrium concentration of ionized oxygen vacancies.

The catalyst productivity is predicted to be proportional to the copper-oxide contact area. The productivity of a series of copper/thoria catalysts is found to be directly proportional to the amount of highly dispersed (XRD-invisible) copper in the sample (Table 2). This is directly related to the metal-oxide contact area as the free copper surface area is negligible.

A critical feature of the theory is the prediction that the response to changes in CO_2/CO reactant ratio depends on the value of $Hv - 2\phi(b)$. This depends on both the metal and the oxide and not on either alone. Furthermore, the mass action treatment above shows that the response to oxygen partial pressure

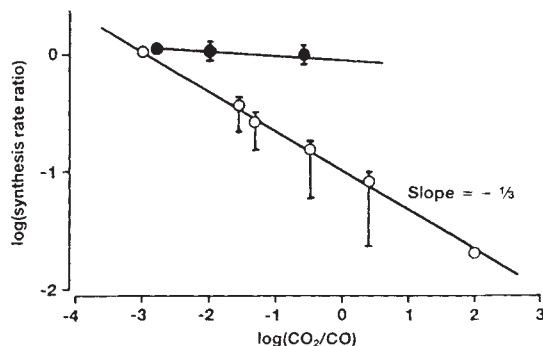


Fig. 4 Experimental synthesis rates as a function of CO_2/CO feed ratio. Open circles, Cu/ThO₂ catalyst; closed circles, Pt/ThO₂ catalyst.

ure will be to the $-1/6$ power for simple oxides in which doubly ionized oxygen vacancies are the dominant defect.

The experimental relative productivity ratios have been determined as follows. The reaction temperature and pressure were typically 473 K and 10 bar. The microreactor consists of a stainless-steel tube, 7-mm internal diameter with an internal thermowell. The catalyst volumes and space velocity used were 1 cm^{-3} and $3,600 \text{ h}^{-1}$. Gas flows were controlled by thermal mass flow controllers and gas analysis achieved by an online gas chromatograph. A catalyst sample was run under a CO_2 -free feed until steady methanol synthesis production (or at least only a slow rate of change) was achieved. The feed was then changed to one with a well defined CO_2/CO ratio, and once again a stable production obtained. Finally, the feed was changed back to a CO_2 -free stream (see Fig. 2). The quantity required is the ratio of the catalyst productivity in CO/H_2 at the oxygen partial pressures of the CO_2 -free and CO_2 -containing feeds. The productivity in the CO_2 -containing stream cannot be used directly to obtain this ratio as CO_2 reacts to form methanol at a much faster rate than CO. But the productivity ratio just before introduction of the CO_2 -containing feed and just after reintroduction of the CO_2 -free feed is just as satisfactory because the solid-state equilibria require a relatively long time (1–2 h) to become re-established.

Figure 3 shows the experimental productivity ratios plotted against metal workfunction for a series of metal/thoria catalysts. The high workfunction metals induce activities that do not depend strongly on the carbon dioxide treatment, whereas the low workfunction metals are strongly deactivated. Figure 4 shows that the productivity ratio becomes proportional to $-1/3 \log(\text{CO}_2/\text{CO})$ for copper/thoria and relatively independent of CO_2/CO for platinum/thoria. Comparison of the experimental productivity ratios for Cu/Zn, Cu/Zr and Cu/Th catalysts (0.8, 4, 10) with the normalized free enthalpy of formation of the oxide (-0.86 , -1.36 , -1.52 eV) shows that the difficult-to-reduce oxides are strongly deactivated by carbon dioxide, as expected. Clearly the response of methanol synthesis catalysts to the oxygen partial pressure depends on the combination of the metal and the oxide, and not on either the metal or the oxide alone.

Table 2 Relative productivities of copper/thoria catalysts as a function of the molar ratio of sub 15 Å copper particles to thoria in the catalyst, as determined by quantitative X-ray diffraction

Relative productivity	Molar ratio Cu/ThO ₂
1.000	1.18
0.804	1.08
0.881	1.00
0.583	0.67
0.327	0.52
0.167	0.25
0.065	0.10

Function of oxygen vacancies

The function of the oxygen vacancies may be understood by combining the junction effects with an actinide organometallic mechanism for CO activation²¹. The initial thoria vacancy is denuded of electrons by the adjacent metal. Hydrogen dissociates at this site and forms a hydride by reclaiming an electron from the solid. The formation of the hydride-vacancy association perturbs the defect equilibria and may lead to substantial hydrogen incorporation into the oxide. Insertion of CO can occur²¹ to give a partially hydrogenated species. Further hydro-

generation is facile until the methoxide ion is formed. The final desorption of a neutral methanol molecule is difficult because the electrons associated with the ion must remain in the solid. With unpromoted oxides, these electrons are left in the high-energy oxygen vacancy site. For the junction-effect-promoted system, either the lower energy metal Fermi level is available or (as the high concentration of surface defects provides good conduction routes) concerted chemisorption of reactants at another vacancy site provides an even more willing acceptor state and may be necessary for desorption to occur. If CO₂ rather than CO insertion occurs, the final step is easier. The extra oxygen can reoxidize the vacancy and the electron released on product desorption can now be accommodated with advantage in the oxide valence band. Methanol synthesis from CO₂ is therefore rapid over any catalyst, but the reaction destroys an active site. If steady-state catalyst activity is to be maintained, the site must be reformed and this process depends on the factors described above.

I do not have space here to discuss in detail the proposal that the copper surface is solely responsible for methanol synthesis,

but we also have found only low methanol synthesis productivity over pure copper catalysts (derived from copper carbonate and supported on carbon)⁶. The correlation between copper surface area and catalyst productivity^{1,22} has been disputed¹³ and certainly large increases in rate (ninefold) are found when zinc oxide is present²³. Nevertheless, some methanol synthesis activity over partially oxidized copper metal cannot be excluded in the limited case of synthesis from CO₂-containing feeds over copper catalysts with significant free copper surface area^{2,22,23}. The attraction of this junction effect theory is that it provides a unified view of all the different methanol synthesis catalysts. Not only have new catalysts been successfully predicted, but the approach provides successful quantitative predictions for the productivity of a range of metals and oxides in CO/H₂ feeds in equilibrium with various oxygen partial pressures.

The author wishes to thank colleagues at the BP Research Centre at Sunbury (particularly Dr P. Meehan and Dr A. G. Mitchell), Professor R. W. Joyner and Professor J. B. Pendry FRS for their help in this work, and BP Research International PLC for permission to publish this paper.

Received 15 April; accepted 6 June 1988.

- Chinchen, G. C. et al. *Am. chem. Soc. Div. Fuel Chem. Symp. on Methanol Synthesis Philadelphia* **29**, 178-185 (1984).
- Chinchen, G. C., Spencer, M. S., Waugh, K. C. & Whan, D. A. *Frady Symp. chem. Soc.* **21**, paper 18 (1986).
- Chinchen, G. C. & Waugh, K. C. *J. Catal.* **97**, 280-283 (1986).
- Klier, K. *Appl. surf. Sci.* **19**, 267-297 (1984).
- Bulko, J. B., Herman, R. G., Klier, K. & Simmons, G. W. *J. phys. Chem.* **83**, 3118-3122 (1979).
- Klier, K. in *Advances in Catalysis* Vol. 31 (eds Eley, D. D., Pines, H. & Weisz, P. B.) 243-310 (Academic, New York, 1982).
- Okamoto, Y., Fukino, K., Imanaka, T. & Teranishi, S. *J. phys. Chem.* **87**, 3747-3754 (1983).
- Fleisch, T. H. & Mievill, R. L. *J. Catal.* **90**, 165-172 (1984).
- Fleisch, T. H. & Mievill, R. L. *J. Catal.* **97**, 284-285 (1986).
- Clausen, B. S., Lengeler, B., Rasmussen, B. S. & Topsoe, H. *Proc. Intl Congr. Prog. X-ray Studies by Synchrotron Radiation Strasbourg* paper 2 (1985).

- Kung, H. H. *Catal. Rev. Sci. Eng.* **22**, 235-259 (1980).
- Friedrich, J. B., Wainwright, M. S. & Young, D. J. *J. Catal.* **80**, 1-24 (1983).
- Bart, J. C. & Sneed, R. P. A. *Catalysis Today* **2**, 1-115 (1987).
- Schwab, G. M. in *Advances in Catalysis* Vol. 27 (eds Eley, D. D., Pines, H. & Weisz, P. B.) 1-22 (Academic, New York, 1978).
- Cox, D. M., Reichmann, K. C., Trevor, D. J. & Kaldor, A. *J. chem. Phys.* **88**, 111-119 (1988).
- Mahan, G. D. *J. appl. Phys.* **54**, 3825-3832 (1983).
- Kroger, F. A. in *The Chemistry of Imperfect Crystals* Vol. 2 (North-Holland, 1974).
- Sze, S. M. in *Physics of Semiconductor Devices* 2nd edn (Wiley, New York, 1981).
- Brillson, L. J. *Phys. Rev. Lett.* **40**, 260-263 (1978).
- Nix, R. M., Rayment, T., Lambert, R. M., Jennings, J. R. & Owen, G. J. *Catal* **106**, 216-234 (1987).
- Mark, T. J. *Science* **217**, 989-997 (1982).
- Chinchen, G. C., Waugh, K. C. & Whan, D. A. *Appl. Catal.* **25**, 101-107 (1986).
- Bridgewater, A. J., Wainwright, M. S. & Young, D. J. *Appl. Catal.* **28**, 241-253 (1986).

Autoregulated instability of β -tubulin mRNAs by recognition of the nascent amino terminus of β -tubulin

Tim J. Yen, Paula S. Machlin & Don W. Cleveland

Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

Tubulin synthesis in animal cells is controlled by an autoregulatory mechanism that modulates the stability of polysome-bound tubulin messenger RNAs. The β -tubulin RNAs are selectively targeted as substrates for destabilization not through the recognition of specific RNA sequences, but rather through co-translational recognition of the amino-terminal β -tubulin tetrapeptide after its emergence from the ribosome. This motif is likely to be used in other systems where RNA degradation is coupled to ribosome attachment and translation.

EXPRESSION of α - and β -tubulins, the two principal constituents of microtubules, is regulated in higher eukaryotes by molecular processes that operate at two levels. The first of these is the transcriptional activation during cell differentiation of one or more members of the small multigene families comprised of about 6 or 7 functional genes that encode either subunit (see ref. 1 for review). In most animal cells, a second regulatory event establishes the appropriate quantitative level of tubulin expression through an autoregulatory pathway in which the apparent intracellular concentration of tubulin heterodimers (comprised of one α - and one β -tubulin polypeptide) modulates the stability of tubulin messenger RNAs²⁻⁷.

The determination of the molecular mechanism through which autoregulated destabilization of tubulin mRNAs is achieved is of general interest for two reasons. First, although autoregulation is well known in prokaryotes, only a few examples have been

found in eukaryotes. Tubulin provides the best-known example of how autoregulation can be achieved for normal cellular genes in higher organisms. Second, it is becoming increasingly apparent that for many genes (for example, histones^{8,9} and a series of lymphokine, cytokine and proto-oncogene mRNAs¹⁰⁻¹⁵), the level of expression is strongly dependent not only on the rate of gene transcription but also on the stability of the mature mRNAs. Despite this, essentially nothing is known of how mRNA half-life is established for almost all eukaryotic mRNAs.

As part of our goal to elucidate the mechanism of post-transcriptional regulation of tubulin gene expression, we have recently determined that the first 13 translated nucleotides of a β -tubulin mRNA (which encode the first four amino acids, Met-Arg-Glu-Ile, of β -tubulin) are sufficient to confer autoregulated instability onto heterologous RNAs if present as the first

translated nucleotides. In addition, as deletions within this domain (either codon 2 or codons 3 and 4) of an authentic β -tubulin gene yield RNAs insensitive to autoregulation, we have concluded this 13-nucleotide segment is both necessary and sufficient to specify RNAs as substrates for autoregulated degradation⁷. Furthermore, using various inhibitors of protein synthesis¹⁶ and by transfection of β -tubulin genes with premature translation termination codons^{7,16}, we have shown that for β -tubulin mRNAs to be substrates for autoregulation, they must be associated with polysomes and that translation of them must proceed past 41 codons.

Taken together, these data suggest two possible models for the mechanism that establishes autoregulated instability of polyribosome-bound β -tubulin RNAs. The first model proposes that RNA degradation is facilitated by unpolymerized tubulin subunits binding (either directly or indirectly) to the 13-nucleotide recognition sequence present in β -tubulin mRNAs. That only ribosome-bound RNAs are substrates for degradation can be explained if translation is required to make the recognition domain accessible for binding. The second model arises from the realization that the sequence conferring autoregulated instability actually encodes the four amino-terminal β -tubulin amino acids, which suggests that the true recognition event is a protein-protein interaction in which free subunits interact (directly or indirectly) with the nascent tubulin polypeptide immediately after it emerges from the large ribosomal subunit. This putative binding event must somehow be transduced through the adjacent ribosome to signal the degradation of the RNA being translated.

To distinguish between these two models (recognition of the nucleotide sequence or recognition of the encoded polypeptide), we have systematically introduced 25 different nucleotide-base substitutions into the 13-base regulatory element of a human β -tubulin gene¹⁷. Some of these mutations alter the amino acids encoded by codons 2 or 3, whereas others leave the polypeptide unchanged (because of the redundancy of the genetic code). We have also changed the translational reading frame of the 13-nucleotide 'identifier' sequence such that its position within the mRNA remains virtually unchanged but it is now translated in an inappropriate reading frame (thereby producing a polypeptide with a different amino-acid sequence). We determined whether RNAs derived from each of these mutant genes remain substrates for autoregulated instability after their re-introduction into animal cells by DNA transfection. Our results show unambiguously that the initial step in the pathway for autoregulated degradation of polysome-bound β -tubulin mRNAs is the recognition by the autoregulatory factor(s) of the nascent, amino terminus (Met-Arg-Glu-Ile) of β -tubulin.

Any Arg codon at aa2 maintains autoregulation

To determine whether autoregulated tubulin mRNA instability is specified by recognition of a nucleotide sequence or of the encoded polypeptide, we used site-directed mutagenesis¹⁸ to introduce one, two or three base substitutions within the 13-nucleotide identifier sequence of the human M40 β -tubulin gene¹⁷. Initially, we introduced all six possible single base substitutions at the first and third positions of the second codon (AGG). We also introduced nine double and one triple base substitution. Each mutant β -tubulin gene contains all the flanking and intervening sequences necessary for expression following transient transfection into mouse L cells. We then assayed cytoplasmic RNAs isolated from parallel dishes of transfected cells that contain either a normal level of unassembled tubulin subunits or an elevated concentration of free subunits (as a consequence of colchicine-induced microtubule depolymerization) to determine the concentration of mutant β -tubulin transcripts and of the endogenous mouse m β 5-tubulin RNA (Fig. 1).

As previously reported¹⁹, RNAs transcribed from the wild-type M40 gene (which carries codon AGG (encoding arginine) at the second amino-acid position) are destabilized (about five- to sixfold) as a consequence of colchicine-induced elevation of

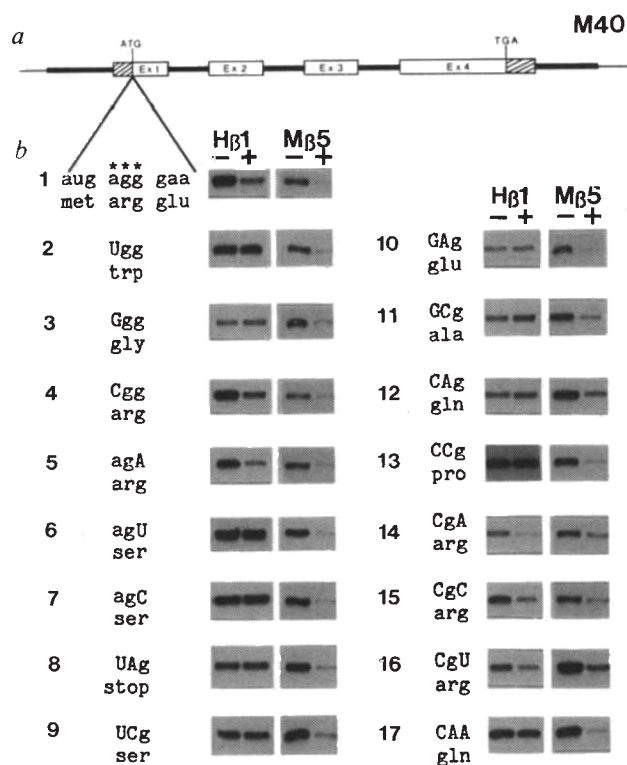


Fig. 1 Autoregulated instability of β -tubulin RNAs with mutations at codon 2. *a*, Schematic diagram of the human β -tubulin gene M40¹⁷, for which we used site-directed mutagenesis to substitute nucleotides at codon 2. Fine line, pUC sequences; bold line, M40 5' and 3' or intron sequences; open boxes, exon (Ex) sequences corresponding to coding regions; hatched boxes, exon sequences corresponding to 5' and 3' untranslated regions. *b*, Quantitative analysis of levels of cytoplasmic RNAs extracted from mouse Ltk⁻ cells transiently transfected with various mutant β -tubulin genes. Wild-type sequence of the first three codons is shown, followed by the various nucleotide substitutions (capitals) introduced at codon 2. Encoded amino acids are indicated below the nucleic acid sequences. Equal amounts of cytoplasmic RNA from cells with normal (-) or elevated (+) unpolymerized tubulin subunits were analysed for m β 5 (an endogenous mouse β -tubulin mRNA) and for RNAs from each transfected tubulin gene.

Methods. Mixed oligonucleotides (19–21mers) were synthesized using an Applied Biosystems DNA Synthesizer and purified by HPLC. Site-directed mutagenesis was performed as described¹⁸ with the following modifications: T4 DNA polymerase was substituted for Klenow and the *in vitro* synthesis reaction was performed at 37 °C for 90 min followed by digestion with *Eco*RI and *Hind*III to release a 900-base-pair tubulin insert containing the DNA modification. This heteroduplex fragment was isolated from low-melting agarose, ligated into M13mp19 and transformed into *Escherichia coli*. Positive plaques were identified by hybridization to ³²P-labelled oligonucleotides and confirmed by sequencing the DNA²². The double-stranded form of the inserts containing the desired mutations were isolated from the recombinant M13 replicative form DNAs and recloned into the human tubulin gene M40 by insertion between the *Hind*III and *Nhe*I sites. Transient transfections using DEAE-dextran were performed as described⁷. Unpolymerized tubulin subunit concentrations were elevated in some samples (+) by treatment of transfected cells with 10 μ M colchicine for the final 5 h of culture. RNA levels determined by S1 analysis⁷.

the tubulin subunit levels. Except for mutant codons CGG and AGA (both of which are alternative arginine codons), the remaining four single-base substitution mutants at codon 2, UGG (Trp), GGG (Gly), AGU (Ser), and AGC (Ser), yield RNAs that are insensitive to autoregulated instability (see rows 4, 5 and 2, 3, 5, 7, respectively, of Fig. 1). Similarly, analysis of

10 additional mutants (nine 2-base and one 3-base substitution) reveals that RNAs from only three of these, CGA (Arg), CGC (Arg), and CGU (Arg), remain substrates for regulated RNA instability (see rows 14, 15, 16). Moreover, for each of the mutants in which autoregulation is still maintained, the sensitivity of each RNA to autoregulated degradation is quantitatively retained (determined by densitometry of the appropriate autoradiographs). In all cases, the autoregulated instability of endogenous β 5-tubulin RNA demonstrates retention of the autoregulatory response in all transfected cells, whereas levels of RNAs encoding ribosomal protein L16 are completely insensitive to colchicine-induced microtubule depolymerization, thus ensuring that comparable levels of RNA are assayed in each sample (data not shown).

Inspection of the nucleotide substitutions of the five mutants that retain autoregulated RNA instability and of the 11 whose RNAs are not regulated correctly does not identify an obvious consensus nucleotide sequence that specifies autoregulation. For example, two double mutants of the initial AGG codon (CGC and CGU) restore autoregulation to single base substitutions (AGC and AGU) that disrupt it (compare rows 15, 16 with rows 6, 7, respectively, of Fig. 1). The most notable feature among the five mutant RNAs which remain substrates for autoregulation is that, like the wild-type gene, all of them encode arginine at codon 2. In fact, these five mutants and the wild type comprise the entire repertoire of arginine codons.

Mutagenesis at codon 3

The mutagenesis at codon 2 clearly supports a model for tubulin RNA instability in which the encoded tubulin polypeptide is the element that targets an RNA as a substrate. To test this hypothesis further, we mutagenized the next three nucleotides of the 13-nucleotide sequence that is necessary and sufficient to confer autoregulation. We constructed nine mutants of the GAA triplet that normally encode glutamic acid at the third residue of β -tubulin and we tested RNAs transcribed from each for autoregulated instability. The results (Fig. 2) show that three of the nine mutant RNAs (with codons GAG (Glu), GAU (Asp) and GAC (Asp)) are still substrates for regulated instability, whereas the remaining six mutants, AAG (Lys), ACC (Asn), CAG (Gln), UAU (Tyr), UAG (stop) and UAC (tyr) are insensitive (see rows 2–4 and 5–10, respectively, of Fig. 2). Comparison of the nucleotide sequences of the nine mutants shows that whereas all four nucleotides are permitted in the third position, additional base substitutions introduced at the first position of the codon invariably abolish recognition of the RNA. This observation can be interpreted in two ways. First, if the recognition sequence is at the nucleic-acid level, the ninth nucleotide (the third position in the codon) within the 13-nucleotide domain does not comprise an essential part of the regulatory element. Alternatively, consistent with our previous findings, the mutant RNAs which remain substrates for autoregulation either retain an encoded glutamic acid or the conservative amino-acid substitution, aspartic acid.

Autoregulated mRNAs must encode MREI

Although the results obtained from analysing tubulin RNAs with mutations at codons 2 and 3 strongly support a model in which the nascent tubulin polypeptide is the element recognized by the autoregulatory factor(s), mutations that alter the encoded amino acids by necessity also alter the nucleic-acid sequence. Thus, it remains possible that the loss of recognition of some mutant mRNAs is the consequence of disruption of a nucleotide-recognition sequence rather than the encoded polypeptide. To resolve this ambiguity, we shifted the translational reading frame of the 13-nucleotide identifier sequence but kept it in essentially the same position within the mRNA. To do this, we exploited an existing hybrid gene (MT-M17- tk_{out} ; see ref. 7) comprised of a metallothionein promoter and 5' untranslated region; the first 17 codons from a mouse β -tubulin gene; a six-nucleotide

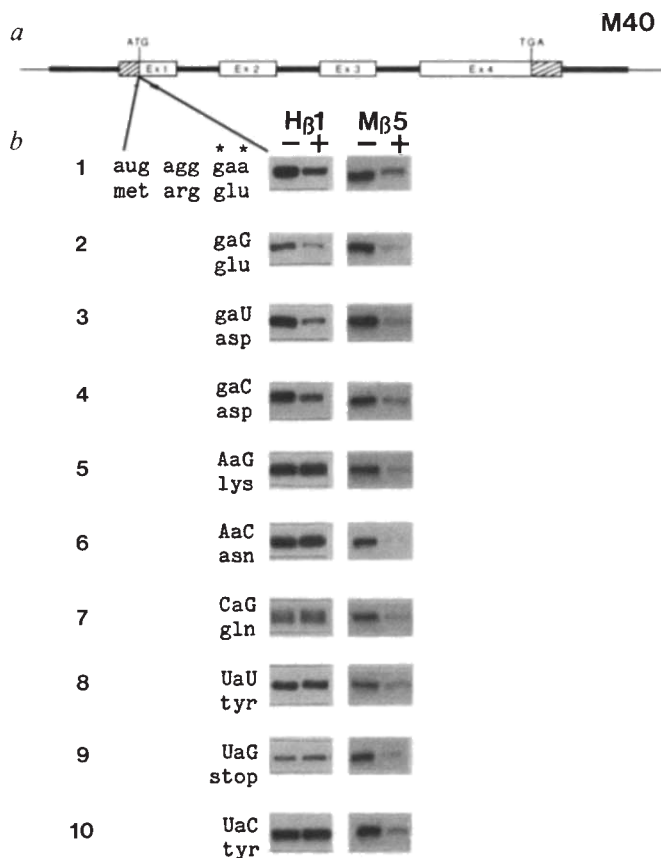
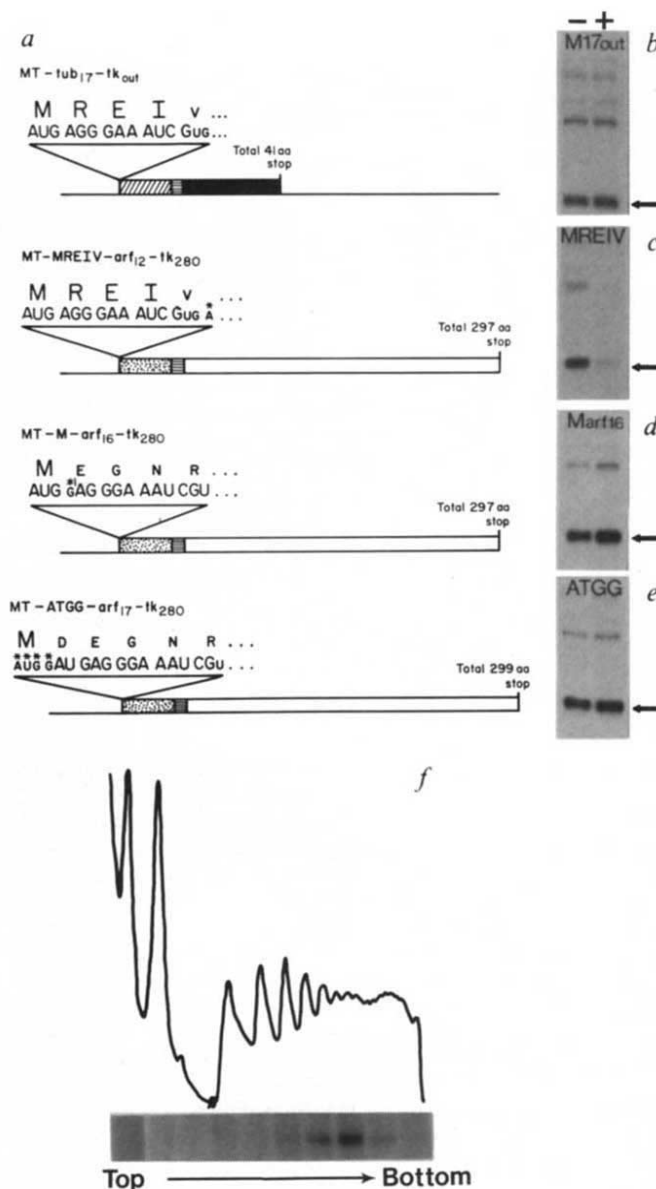


Fig. 2 Autoregulated instability of tubulin RNAs with mutations in the third codon. *a*, Schematic diagram of the human β -tubulin gene M40 and of the mutagenesis at the first and third positions of codon 3. *b*, Quantitative analysis of cytoplasmic RNAs isolated from aliquots of cells transfected with each of nine β -tubulin genes that carry mutations in the third codon (see Fig. 1 for methods).

linker; and 278 codons plus 3' flanking sequences from the herpesvirus thymidine kinase (tk) gene. The translation of this mRNA begins at the authentic tubulin AUG translation initiation codon but the tubulin- tk junction is joined such that translation reads into the wrong reading frame of tk , thereby leading to premature translation termination after the 41st codon (17 amino acids from tubulin, 2 from the linker and 22 from the alternate reading frame of tk). We have previously used this construct to demonstrate that RNAs yielding such premature translation termination products are not substrates for autoregulation, whereas analogous RNAs that are joined in-frame (and yield long translation products) are efficient substrates⁷.

Because some feature of the autoregulatory process requires translation beyond codon 41, autoregulation should be restored to RNAs from MT-M17- tk_{out} by one-base insertions that restore translation to the correct reading frame of tk . We strategically introduced single nucleotide insertions into MT-M17- tk_{out} immediately after the initiating AUG of tubulin or immediately after the fifth tubulin codon. In both cases, the reading frame of tk is restored such that a 297-amino-acid fusion polypeptide is produced. Although the nucleotide sequences of these two hybrid genes are virtually identical, their translation products are different (see Fig. 3*a*). An insertion immediately following the AUG (to produce MT-M- arf_{16} - tk_{280}) will alter the translation reading frame of the remaining 48 tubulin-coding nucleotides to produce a 16-amino-acid polypeptide (encoded by the wrong frame of β -tubulin) which is fused with the remaining 280 amino acids of tk . As a control, insertions introduced after the fifth tubulin codon (MT-MREIV- arf_{12} - tk_{280}) produce a fusion polypeptide comprised of the amino-terminal five β -

Fig. 3 Alteration of the translational reading frame of the β -tubulin RNA sequences abolishes autoregulation. *a*, Schematic diagram of MT-tubulin-*tk* genes into which nucleotide insertions (asterisks) were introduced to restore the proper reading frame to MT-M17-*tk*_{out}. Fine line, untranslated sequences; hatched boxes, tubulin amino acids; open boxes, *tk* amino acids; lined box, two-amino-acid linker; stippled box, amino acids derived from an alternate reading frame of tubulin sequences; closed box, amino acids from an alternative reading frame of *tk*. Nucleotides and amino acids which comprise the regulatory element are represented by large letters. Smaller letters denote sequences (nucleotide or amino acid) not essential for autoregulation. MT, mouse metallothionein promoter and 5' untranslated sequences; M17, mouse β -tubulin encoding the amino-terminal 17 β -tubulin amino acids; *arf*, alternate reading frame of β -tubulin as a result of nucleotide insertions; *tk*, thymidine kinase. The numbers in subscript denote number of amino acids or number of codons. Single letter designations for amino acids are presented. *b-e*, Mouse Ltk⁻ cells were transfected with MT-M17_{out}-*tk* (M17_{out}-*b*), MT-MREIV-*arf*₁₂-*tk*₂₈₀ (MREIV-*c*), MT-M-*arf*₁₆-*tk*₂₈₀ (Marf₁₆-*d*), or MT-ATGG-*arf*₁₈-*tk*₂₈₀ (ATGG-*e*). Equal amounts of cytoplasmic RNA from cells with normal (-) or elevated (+) unpolymerized tubulin subunits were analysed for endogenous m β 5 RNAs and for RNAs from each transfected tubulin gene. (The faint larger species seen in each lane represents RNAs initiated at minor upstream transcriptional initiation sites of the metallothionein (MT) promoter.) Arrows mark the positions for fragments protected by RNAs initiated from the major transcriptional start site in MT. *f*, The distribution of RNAs derived from MT-ATGG-*arf*₁₈-*tk*₂₈₀ within a polysome profile. Polysomes from cells transfected with MT-ATGG-*arf*₁₈-*tk*₂₈₀ were analysed on a 10–40% sucrose gradient¹⁶. The gradient was divided into 10 fractions and RNA isolated from each and analysed by S1 nuclease analysis. Polysome isolation was performed as described¹⁶ with the following modifications: 100-mm diameter dishes of mouse L cells transfected with MT-ATGG-*arf*₁₈-*tk*₂₈₀ were lysed using NP40-lysis buffer (10 mM Tris-HCl, pH 8.6, 100 mM NaCl, 10 mM MgCl₂, 0.5% NP40, 10 mM vanadyl-ribonucleoside complexes and 100 μ g ml⁻¹ of cycloheximide). Gradients (10–40% sucrose) in lysis buffer without NP40 and cycloheximide were centrifuged for 2.5 h at 27,000 r.p.m. in an SW41 rotor¹⁶.



tubulin amino acids (the shortest sequence required for autoregulation contains the first four amino acids) fused to a 12-amino-acid polypeptide derived from the translation of the alternate reading frame (*arf*) of the remaining portion of the 36-nucleotide tubulin-coding sequence which in turn is joined in-frame to *tk*.

If the recognition element is indeed the amino-terminal tubulin peptide, we would predict that RNAs with an insertion altering the reading frame of the amino-terminal tubulin-coding sequence will not be substrates for autoregulation, whereas those RNAs that encode the amino-terminal five tubulin amino acids would be substrates. (Before the introduction of these nucleotide insertions, inspection of the tubulin nucleotide sequence within MT-M17_{out}-*tk* revealed that a translation termination codon would be encountered in the inappropriate tubulin reading frame. To overcome this, we substituted by site-directed mutagenesis the T at position 51 of the tubulin sequence to a C before additional nucleotide insertions were introduced.)

We transfected the constructs MT-M-*arf*₁₆-*tk*₂₈₀ and MT-MREIV-*arf*₁₂-*tk*₂₈₀ into mouse cells and tested their encoded RNAs for autoregulated degradation after colchicine-induced elevation of tubulin subunit concentration. Although transcripts derived from MT-M17_{out}-*tk* are unaffected by increases in tubulin subunit concentrations (Fig. 3*b*), a single base insertion

(A) after the fifth tubulin codon in MT-M17_{out}-*tk* restores autoregulated instability of RNAs transcribed from this gene (MT-MREIV-*arf*₁₂-*tk*₂₈₀; Fig. 3*c*). (Similarly, introduction of a C after the fifth tubulin codon also restored autoregulation; data not shown.) In contrast, transcripts with an insertion positioned immediately after the initiating AUG are not autoregulated (Fig. 3*d*). We eliminated the possibility that the failure of these latter RNAs to be recognized stems from their inability to be translated into a long polypeptide by demonstrating that they are present in the fast-sedimenting region of a polysome profile (data not shown).

Although these results further support the notion that the recognition element for tubulin autoregulation is the amino-terminal peptide, it still remained possible that the single base insertion introduced after the initiating AUG disrupts the putative RNA recognition sequence. To settle this question unambiguously, we constructed one additional mutant gene (MT-ATGG-*arf*₁₇-*tk*₂₈₀), again starting with MT-M17-*tk*_{out}. This time, however, we restored the *tk* sequences to the proper translational reading frame not by insertion of a single nucleotide but rather by a four-base insertion (ATGG) just before the authentic β -tubulin translation initiation codon. In the final gene (Fig. 3*a*), the 13-nucleotide sequence capable of conferring autoregulation remains intact, although its position in the

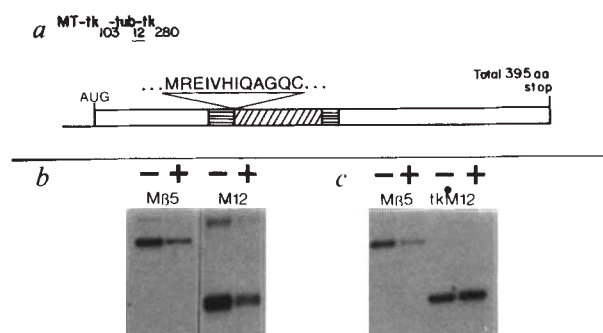


Fig. 4 Translocation of the β -tubulin tetrapeptide recognition sequence within *tk* prevents its recognition by the autoregulatory factor(s). *a*, Schematic diagram of MT-*tk*₁₀₃-*tub*₁₂-*tk*₂₈₀, a construct that internalizes the amino-terminal 12 β -tubulin amino acids within the *tk* polypeptide. The actual encoded polypeptide comprises 97 aa of *tk* followed by the 12 amino-terminal β -tubulin residues, a 6 aa linker and 280 additional aa of *tk*. See legend to Fig. 3a for explanation of symbols. *b*, *c*, S1 nuclease analysis of cytoplasmic RNA isolated from cells transfected with (*b*) MT-*tub*₁₂-*tk*₂₈₀ or (*c*) MT-*tk*₁₀₃-*tub*₁₂-*tk*₂₈₀. M β 5: RNAs from an endogenous mouse β -tubulin gene; M12: RNAs from MT-*tub*₁₂-*tk*₂₈₀; tkM12: RNAs from MT-*tk*₁₀₃-*tub*₁₂-*tk*₂₈₀. Equal amounts of cytoplasmic RNA from cells with normal (–) or elevated (+) unpolymerized tubulin subunits were probed for m β 5 and for RNAs from each transfected gene.

mRNA has been moved four bases. Assuming that translation initiates at the first AUG (that provided by the inserted ATGG), a 299-amino-acid polypeptide will be produced that consists of a methionine followed by 18 amino acids encoded by an alternative translation frame of the tubulin sequence, a 2-amino-acid linker, and the carboxy-terminal 278 amino acids of TK. The levels and polysomal distribution of RNA accumulated in cells transfected with MT-ATGG-*arf*₁₇-*tk* are shown in Fig. 3e, f. Clearly, these RNAs are not substrates for facilitated degradation after elevation of tubulin subunit concentrations. Further, because most MT-ATGG-*arf*₁₇-*tk*₂₈₀ RNAs are in the fast-sedimenting polysomal fractions (Fig. 3f), translation of these mRNAs must initiate at the 5'-most AUG codon, as initiation at the second methionine codon (that specified by the intact 13-base tubulin recognition sequence) can yield only a 41-amino-acid translation unit which is too short to attach simultaneously to more than 2–3 ribosomes⁷. Thus, the disruption of autoregulation in MT-ATGG-*arf*₁₇-*tk*₂₈₀ RNAs must be caused by the failure to translate the β -tubulin amino-terminal tetrapeptide (Met-Arg-Glu-Ile), as the 13-nucleotide RNA sequence is uninterrupted and resides in essentially the same position in the mRNA.

Internalization of MREI blocks recognition

From the cumulative results of the site-directed mutagenesis experiments, we conclude that the recognition of polysome-bound β -tubulin mRNAs by the autoregulatory factor(s) occurs via direct protein-protein interactions involving the nascent β -tubulin polypeptide. To examine whether the position of this

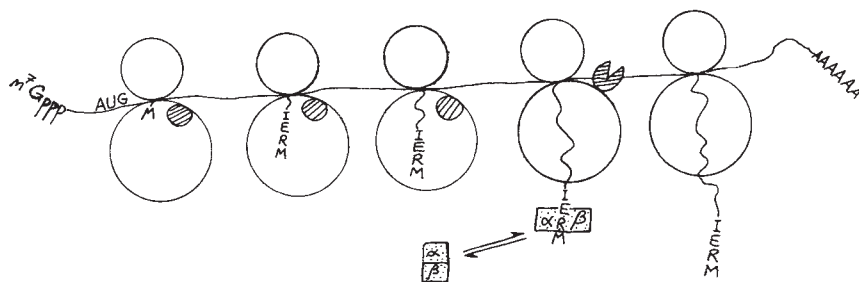
polypeptide sequence relative to the remaining portion of the protein is important for proper recognition by the autoregulatory mechanism, we constructed a gene (derived from MT-M12-*tk*_{in}⁷), in which we placed the amino-terminal 12 tubulin amino acids (in the correct translational reading frame) within a *tk* polypeptide. This construct, MT-*tk*₁₀₃-*tub*₁₂-*tk*₂₈₀ (see Fig. 4), encodes a polypeptide comprised of 395 amino acids (97 from *tk*, 6 encoded by a linker, 12 from the amino terminus of β -tubulin and 280 corresponding to the carboxy terminus of *tk*). We transfected this construct and examined RNAs derived from it for autoregulated instability. As is shown in Fig. 4b, RNAs which are translated into a fusion polypeptide that carries 12 β -tubulin amino acids at its amino terminus are sensitive to autoregulation (M12), whereas those RNAs that encode the identical 12 tubulin amino acids but located within the *tk* polypeptide are insensitive to the tubulin subunit concentration (Fig. 4c). Thus, an autoregulated RNA must encode the amino-terminal β -tubulin amino acids; it must be translated to yield a long polypeptide; and the position of the tubulin residues within the polypeptide is important.

A model for autoregulated mRNA instability

The transfection experiments with gene constructs carrying altered nucleotide sequences and/or translation frames lead us to conclude that the recognition element for autoregulated RNA instability is actually the encoded amino-terminal tetrapeptide (Met-Arg-Glu-Ile). We have derived a model for autoregulated instability from these and from our previous work (Fig. 5). As the unpolymerized tubulin subunit pool is elevated, the initial event that leads ultimately to regulated degradation of β -tubulin RNAs is the interaction between the autoregulatory factor(s), possibly tubulin itself (as shown in Fig. 5) and the nascent amino-terminal tubulin tetrapeptide just after it emerges from the ribosome. How this protein-protein interaction is transduced through the ribosome to yield degradation of the corresponding RNA is not yet known, although there are two general possibilities. First, the binding event could activate a cellular RNase (which itself might be a peripheral ribosomal component). Alternatively, binding could induce a transient stalling of the ribosome that leaves the RNA in an exposed conformation that is a better substrate for non-specific RNases. We cannot yet distinguish between these two possibilities. Further, although we show the binding event occurring immediately after the amino terminus of β -tubulin emerges from the ribosome, the actual position along the polysome profile where this interaction occurs has not been determined experimentally.

That so short a sequence as a tetrapeptide can be specifically recognized may at first seem surprising. But inspection of the predicted amino-terminal four amino acids among all β -tubulin genes sequenced to date reveals that the tetrapeptide Met-Arg-Glu-Ile is absolutely conserved at the amino terminus. Furthermore, a search of the protein database for other proteins that contain this tetrapeptide sequence identifies only β -tubulins. Even considering the possibility that the true minimum sequence is only Met-Arg-Glu (we have not obtained and tested a construct with only the first three amino-terminal β -tubulin codons), a search of the protein-sequence library shows that although

Fig. 5 Proposed model for autoregulated instability of β -tubulin mRNA. Unpolymerized tubulin subunits bind directly (or activate a factor(s) which binds) to the nascent amino-terminal tetrapeptide (Met-Arg-Glu-Ile) of β -tubulin. This binding is transduced through the adjacent ribosome to activate an RNase which degrades the ribosome-bound mRNA. The RNase has been drawn (shaded areas) to be ribosome-associated, although this has not yet been demonstrated. MREI: the amino-terminal β -tubulin polypeptide.



several known proteins carry this tripeptide sequence internally, only tubulins (α and β) have it at their amino termini. Further, as translocation of the first 12 β -tubulin amino acids internally within a *tk* polypeptide disrupts its recognition by the autoregulatory factor(s), the context in which Met-Arg-Glu-Ile appears is important. It seems plausible that an amino-terminal localization may be necessary for autorecognition.

The model easily explains why RNAs that yield only short translation products fail to be autoregulated even though they are efficiently associated with ribosomes⁷. Because it is known for bacterial ribosomes that 30 to 40 amino acids are contained in a tunnel in the large ribosomal subunit²⁰, at least this many amino acids must be translated before the nascent Met-Arg-Glu-Ile peptide would be accessible for binding by the autoregulatory factor(s). Obviously, if translation terminates before this point, it would not be possible to link tubulin RNA levels to subunit concentration. Although we have not formally determined the minimum number of codons through which translation must proceed so that the amino terminus of β -tubulin emerges from the large ribosomal subunit, translation of as few as 90 codons is sufficient for an RNA to be sensitive to autoregulated de-

stabilization (T.J.Y. and D.W.C., unpublished).

In addition to the tubulin example reported here, an increasing number of eukaryotic proteins are regulated at least in part by RNA stability. Examples include many proto-oncogene, lymphokine and cytokine mRNAs that contain within their 3' untranslated regions a ~50-base AU-rich sequence that destabilizes the mRNA¹⁴. Cell-cycle-dependent expression of histone mRNAs is also established in part through changes in mRNA stability specified by sequences carried in the 3'-untranslated region^{9,21}. Although the mechanisms of degradation must differ for each of these RNAs, a striking parallel does emerge: RNA instability is linked to ongoing protein synthesis (tubulin (ref. 16 and this paper); *c-myc*¹⁰; *c-fos*¹¹⁻¹³; and histones⁹). Recognizing this common linkage of mRNA stability to translation, we propose the existence of a widely used motif in which the degradation of RNAs is integrally linked to ribosome attachment and translation. We also make the appealing suggestion that the RNase(s) are localized to actively translating ribosomes.

This work has been supported by a NIH grant to D.W.C. T.J.Y. has been supported by an American Cancer Society Postdoctoral Fellowship.

Received 14 April; accepted 7 June 1988.

1. Cleveland, D. W. *J. Cell Biol.* **104**, 381-383 (1987).
2. Cleveland, D. W., Lopata, M. A., Sherline, P. & Kirschner, M. W. *Cell* **25**, 537-546 (1981).
3. Caron, J. M., Jones, A. L., Rall, L. B. & Kirschner, M. W. *Nature* **317**, 648-650 (1985).
4. Caron, J. M., Jones, A. L. & Kirschner, M. W. *J. Cell Biol.* **101**, 1763-1772 (1985).
5. Pittenger, M. F. & Cleveland, D. W. *J. Cell Biol.* **101**, 1941-1952 (1985).
6. Gay, D. A., Yen, T. J., Lau, J. T. Y. & Cleveland, D. W. *Cell* **50**, 671-679 (1987).
7. Yen, T. J., Gay, D. A., Pachter, J. S. & Cleveland, D. W. *Molec. cell. Biol.* **8**, 1224-1235 (1988).
8. Bird, R. C., Jacobs, F. A. & Sells, B. H. *Biochem. Cell Biol.* **64**, 99-105 (1986).
9. Graves, R. A., Pandey, N. B., Chodchoy, N. & Marzluff, W. F. *Cell* **48**, 795-804 (1987).
10. Linial, M., Gunderson, N. & Groudine, M. *Science* **230**, 1126-1131 (1985).

11. Kruijer, W., Cooper, J. A., Hunter, T. & Verma, I. *Nature* **312**, 711-716 (1984).
12. Miller, A. D., Curran, T. & Verma, I. *Cell* **36**, 51-60 (1984).
13. Muller, R., Bravo, R., Burckhardt, J. & Curran, T. *Nature* **312**, 716-720 (1984).
14. Shaw, G. & Kamen, R. *Cell* **46**, 659-667 (1986).
15. Treisman, R. *Cell* **42**, 889-902 (1985).
16. Pachter, J. S., Yen, T. J. & Cleveland, D. W. *Cell* **51**, 283-292 (1987).
17. Lee, M. G.-S., Lewis, S. A., Wilde, C. D. & Cowan, N. J. *Cell* **33**, 477-486 (1983).
18. Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468-500 (1983).
19. Lau, J. T. Y., Pittenger, M. F. & Cleveland, D. W. *Molec. cell. Biol.* **5**, 1611-1620 (1985).
20. Yonath, A., Leonard, K. R. & Wittman, H. G. *Science* **236**, 813-816 (1987).
21. Capasso, O., Blecker, G. C. & Heintz, N. *EMBO J.* **6**, 1825-1831 (1987).
22. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

Simple RNA enzymes with new and highly specific endoribonuclease activities

Jim Haseloff & Wayne L. Gerlach

CSIRO Division of Plant Industry, Canberra, Australia

In vitro mutagenesis of sequences required for the self-catalysed cleavage of a plant virus satellite RNA has allowed definition of an RNA segment with endoribonuclease activity. General rules have been deduced for the design of new RNA enzymes capable of highly specific RNA cleavage, and have been successfully tested against a new target sequence.

CERTAIN naturally occurring RNA molecules possess the property of self-catalysed cleavage¹. One class of this reaction is shared by a number of small circular RNA molecules which replicate in plants, either alone (viroid RNAs) or dependent on a helper virus (satellite RNAs). Self-cleavage has been demonstrated *in vitro* for avocado sunblotch viroid (ASBV)² and the satellite RNAs of tobacco ringspot virus (sTobRV)^{3,4} and lucerne transient streak virus (sLTSV)⁵ and appears to be an essential and unique part of the life cycle of these RNAs. During replication, circular forms of the RNAs are thought to act as templates for the production of longer than unit-length RNA transcripts⁶⁻⁸, and the subsequent self-catalysed cleavage of these concatameric transcripts gives rise to unit-length progeny. In addition to these replicating RNAs, *in vitro* self-catalysed cleavage has now been observed in RNA transcripts of certain tandemly repeated DNA sequences from newt⁹.

These self-catalysed RNA cleavage reactions share a requirement for divalent metal ions and neutral or higher pH, and result in the production of RNA with termini possessing 5', hydroxyl and 2', 3' cyclic phosphate groups^{3,5,9,10}. The cleavage

reactions presumably result from RNA conformation bringing reactive groups into close proximity. The sites of cleavage are specific and associated with domains of conserved sequence and secondary structure. For sTobRV (ref. 4 and J.H. and W.L.G., unpublished results), ASBV¹¹ and sLTSV¹², it is known that precisely these conserved regions are required for cleavage, and may thus be directly involved in the reaction. A consensus of the domains associated with known RNA self-cleavage reactions is shown schematically in Fig. 1a. They consist of three branched RNA helices which flank the susceptible phosphodiester bond and two single-strand regions which are highly conserved in sequence. The helices are numbered I, II and III, following the convention of Forster and Symons⁵, but the structures are redrawn with the sequences immediately adjacent to the site of cleavage shown across the top of the diagram. In the different self-cleavage domains, the base-paired stems are variously terminated by a small single-strand loop, or connected to the remainder of the viroid or satellite RNA molecule.

The self-catalysed cleavage of these RNAs is normally an intra-molecular reaction, that is a single molecule contains all

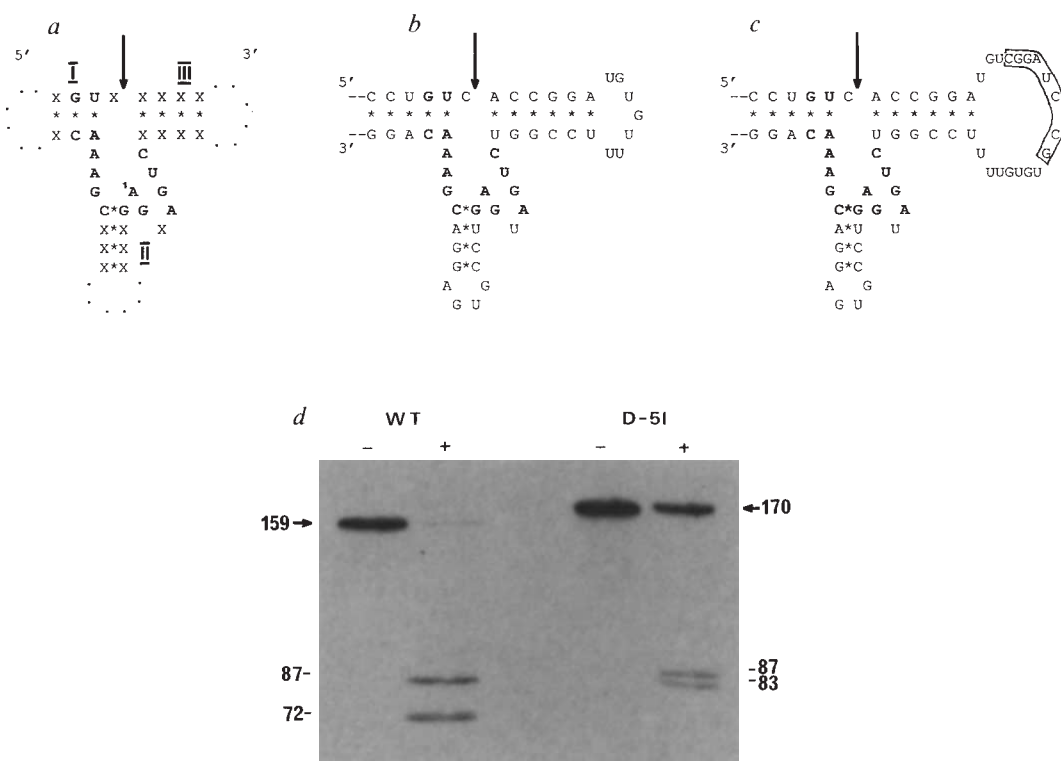


Fig. 1 A mutated self-cleavage domain which is active. *a*, The conserved structures associated with naturally-occurring RNA cleavage sites in ASBV^{2,11}, newt satellite DNA transcripts⁹ and the satellite RNAs of TobRV⁴, LTSV⁵, velvet tobacco mottle virus²², *Solanum nodiflorum* mottle virus²² and subterranean clover mottle virus²³ are summarized. Nucleotide sequences which are absolutely conserved between these structures are shown, while others are represented by X. The three RNA helices (I, II and III) vary in length from 2 to 7 base pairs. Base-pairing is represented by (*) and the site for RNA cleavage is arrowed. (1 an extra U is positioned after this residue in LTSV (+) strand.) *b*, The conserved 52-nucleotide structure associated with self-cleavage of (+) strand sTobRV RNA is shown schematically. We refer to this as the wild-type cleavage structure. Conserved residues are highlighted, and the site for cleavage is arrowed. *c*, The D-51 *in vitro* mutant of sTobRV possesses an altered self-cleavage domain, containing an insertion of 8-nucleotides (CGGAUCCG, shown boxed) together with a flanking duplication of three nucleotides (UGU, 7-9). *d*, Subcloned *Hae*III fragments of wild-type sTobRV and the D-51 *in vitro* mutant were each transcribed in both (-) and (+) orientations and radiolabelled transcripts were fractionated by PAGE. The positions of uncleaved 159 and 170 base transcripts from the wild-type (WT) and mutant (D-51) sequences are arrowed; sizes of cleavage products are shown.

Methods. 97 and 108 base-pair *Hae*III fragments containing the sites for self-cleavage were excised from sequenced plasmid clones containing wild-type and D-51 sTobRV sequences, respectively. The fragments were each ligated into the *Sma*I site of pGEM 4 and screened to obtain both orientations of the insert. The plasmids were linearized using *Eco*RI, and (+) and (-) strand RNAs of lengths 159 and 170 bases were transcribed using 200 units ml⁻¹ T7 RNA polymerase in 50 mM Tris-HCl pH 7.5, 10 mM NaCl, 6 mM MgCl₂, 2 mM spermidine, 1,000 unit ml⁻¹ RNasin, 500 μ M ATP, CTP and GTP with 200 μ M [α -³²P]UTP. RNAs were fractionated by 10% polyacrylamide, 7 M urea, 25% formamide gel electrophoresis, and autoradiographed.

the RNA-encoded functions required for cleavage. For example, cleavage of sTobRV (J.H. and W.L.G., unpublished results) and sLTSV RNAs¹² requires single contiguous sequences of about 52 nucleotides in length, which contain the conserved structures shown in Fig. 1a. The self-cleavage of ASBV in contrast requires the participation of two sequences which are widely separated on the intact RNA. Similar sequences have been transcribed *in vitro* as separate 19- and 24-nucleotide RNA fragments¹¹. When mixed, the two fragments may base-pair (via stems I and II) to form the conserved structure associated with self-cleavage (Fig. 1a) and efficient cleavage of the 24-nucleotide RNA was observed. The 19-nucleotide fragment remained unaltered and could participate in many cleavage reactions, and therefore possessed the properties of an RNA enzyme. In this case, however, the cleaved RNA also contained conserved sequences and secondary structure, making this system unsuitable as a general model for the design of ribonucleolytic RNA enzymes with wide sequence specificity. Similarly, the more complex RNA component of RNase P¹³ and shortened forms of the self-splicing ribosomal RNA intervening sequence from *Tetrahymena thermophila*^{14,15}, which have both been shown to act as endoribonucleases, have so far proved of limited practical use for the design of new endonuclease activities. The term

'ribozyme' has been used to describe such RNAs with catalytic activity¹⁵.

We have taken the single self-cleaving domain from the (+) strand of sTobRV and dissected its RNA substrate and enzyme activities. An RNA substrate was defined which possessed little conserved sequence and no essential secondary structure. Inspection of the separated substrate and ribozyme activities, and comparison with other naturally occurring self-cleaving domains, led to a model for the design of oligoribonucleotides which possess new and highly sequence-specific endoribonuclease activities. This model was successfully tested by the design and construction of ribozymes targeted against three sites within the Tn9 chloramphenicol acetyl-transferase (CAT) messenger RNA sequence.

A mutant that self-cleaves

Encapsidated forms of the satellite RNA of tobacco ringspot virus (+strand sTobRV) have been shown to undergo self-catalysed cleavage at a specific phosphodiester bond³ between bases 359 and 1. RNA transcripts of sTobRV made *in vitro* also undergo efficient self-cleavage⁴, and *in vitro* mutagenesis and transcription of cloned sTobRV complementary DNA has led to mapping of the sequences required for the reaction (J.H. and

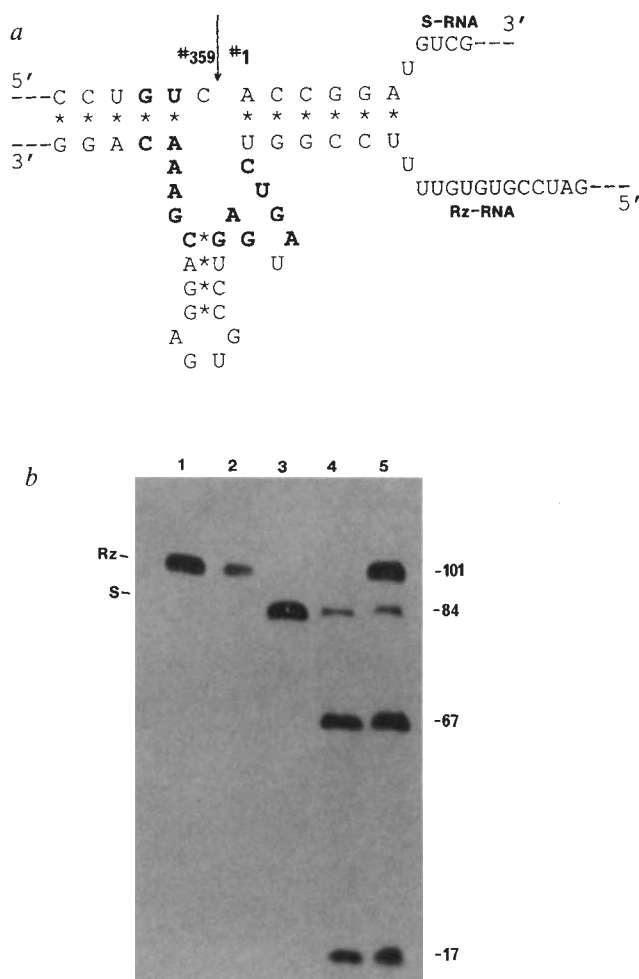


Fig. 2 Separation of substrate and catalytic activities. *a*, The inserted nucleotides in the D-51 mutant (Fig. 1c) contain a *Bam*HI restriction endonuclease site. *Bam*HI was used to split the mutant DNA, and the two sequences were subcloned and transcribed separately *in vitro*. The RNA transcripts are shown schematically, with potential base-pairings between the RNAs indicated (*). The fragment containing the arrowed site for cleavage was termed S-RNA, and the other Rz-RNA. *b*, [32 P]Rz transcript (101 bases) was incubated alone (lane 1), and with unlabelled S-RNA (lane 2). [32 P]S RNA was incubated alone (lane 3), and with unlabelled and [32 P]Rz RNAs (lanes 4 and 5, respectively). The sizes of the two products are consistent with cleavage of the S-RNA (84 bases) at the normal site between nucleotides 359 and 1, to give 5' and 3' proximal fragments of 67 and 17 nucleotides, respectively. **Methods.** Isolated D-51 *Hae*III fragment was digested with *Bam*HI and the Rz and S fragments ligated into *Sma*I-*Bam*HI digested pGEM3 and 4, respectively. T7 DNA polymerase treated, *Kpn*I-digested Rz-pGEM3 and *Xba*I digested S-pGEM4 were transcribed using SP6 RNA polymerase under the same conditions used for T7 RNA polymerase (Fig. 1), either without radiolabel or in the presence of 200 μ M [α - 32 P]UTP. The Rz and S RNAs were incubated alone in 50 mM Tris-HCl pH 8.0, 20 mM MgCl₂ at 50 °C for 60 min, fractionated on a 10% polyacrylamide, 7 M urea, 25% formamide gel and autoradiographed.

W.L.G., unpublished results). In these experiments, cloned sTobRV cDNAs were mutagenized using an oligonucleotide linker insertion protocol. A library of sTobRV mutants resulted, and nucleotide sequence analysis showed that each mutant contained an inserted *Bam*HI linker sequence (CGGATCCG) together with flanking duplicated or deleted sTobRV sequences. The mutants were transcribed *in vitro* and the RNAs assayed for their ability to undergo cleavage. The only mutations affect-

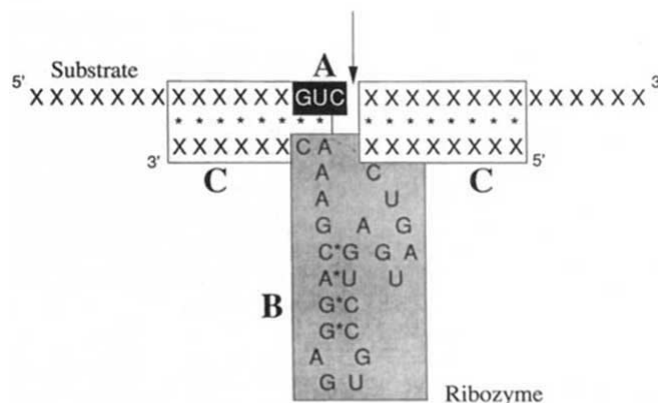


Fig. 3 Model for design of ribozymes. Three structural domains are boxed. (A) contains conserved sequences in the RNA substrate, immediately adjacent to the site of cleavage. (B) comprises the highly conserved sequences maintained in the ribozyme, and the regions (C) consist of flanking helices, with base-pairing between the substrate and ribozyme RNAs.

ing cleavage of (+) strand sTobRV were those which altered a 12-nucleotide sequence adjacent to the cleaved bond. This sequence contains the domain of conserved sequence and possible secondary structure required for self-cleavage in other RNAs (Fig. 1a,b).

One mutant, designated D-51, which contained an alteration in this domain proved an exception (Fig. 1c). It contained an 8-nucleotide *Bam*HI linker sequence inserted between three duplicated sTobRV nucleotides numbered 7 to 9, yet still underwent cleavage. The wild-type and D-51 cleavage domains were subcloned as 97 and 108 base-pair *Hae*III fragments, respectively, and RNAs were transcribed in both orientations *in vitro*. As expected, no cleavage of (-) strand RNA transcripts was observed, but (+) strands of both the wild-type and D-51 sequences underwent cleavage (Fig. 1d), with cleavage of the D-51 RNA being somewhat less efficient than that of the wild-type. The extra 11-nucleotides contained in the D-51 mutant were precisely located within a proposed single-strand loop in the conserved structural domain associated with self-cleavage (Fig. 1c). This provided further experimental support for the structural model, and suggested that this single-strand loop region may not be required for cleavage.

Separation of substrate and nuclease activities

Using the *Bam*HI restriction endonuclease site inserted into D-51, flanking *Hae*III-*Bam*HI and *Bam*HI-*Hae*III fragments were obtained and each was subcloned into an *Escherichia coli* plasmid suitable for *in vitro* RNA transcription. This allowed us to effectively eliminate the mutated single-strand loop from the self-cleavage domain, splitting the region into two RNA segments (Fig. 2a). The smaller *Hae*III-*Bam*HI transcript contained sTobRV nucleotides 321 to 9, including the actual site of cleavage, and was termed the substrate or S-RNA. The *Bam*HI-*Hae*III transcript containing sTobRV nucleotides 7 to 48 was termed the ribozyme or Rz-fragment. These were separately transcribed with and without radiolabelled ribonucleotide as tracer. Both the S- and Rz-RNAs showed no significant degradation when incubated alone (Fig. 2b, lanes 1 and 3) under conditions suitable for highly efficient self-cleavage (50 °C, 20 mM MgCl₂, pH 8.0). The Rz-RNA also appeared unaltered after incubation with the S-RNA (Fig. 2b, lanes 2 and 5), but efficient cleavage of the S-RNA occurred (Fig. 2b, lanes 4 and 5) producing two fragments. The product sizes were consistent with cleavage at the normal site. These data show that the S-RNA acted as a substrate for ribonucleolytic cleavage by the Rz-RNA, which apparently acted in a catalytic fashion.

The substrate defined above contains none of the highly conserved secondary structures and few of the conserved sequences thought to be essential for cleavage. Instead, base-pairing between the sTobRV enzyme and substrate RNA segments (via stems I and III) may result in formation of an RNA complex with the required properties for cleavage. Inspection of the separated substrate and enzyme activities derived from the sTobRV (+) strand, and comparison with other naturally occurring self-cleavage sites has led to a model for the design and construction of novel ribozymes.

Design of new ribozymes

A model showing the proposed minimal structural requirements for ribozyme-catalysed RNA cleavage is presented in Fig. 3, and consists of three elements. First, a region (A) containing the sequence GUC adjacent to the site for cleavage is brought into close proximity to a second region (B) of highly conserved sequence and secondary structure. Third, flanking regions (C) of base-paired RNA helix stabilize this interaction. The model provides the basis for the *de novo* design of ribozymes with new sequence-specific endoribonuclease activities by consideration of three main parameters.

(1) Before a ribozyme can be designed, a particular site for cleavage must be chosen within the target RNA. In naturally occurring self-cleavage, the site in the RNA substrate is 5' flanked by several nucleotides which are highly conserved (A). This sequence, GUC, usually immediately precedes the site of cleavage. In one exception, GUA precedes the active site of cleavage in the (−) strand of sLTSV⁵. Alteration of a self-cleaving sequence similar to that found in transcripts of newt satellite DNA also showed that the residues immediately preceding the cleavage site could be changed¹⁶. When the normal sequence GUC was changed to GUA or GUU, cleavage activity remained essentially unaltered, but no cleavage occurred after the sequence GUG. Further, cleavage was shown after the sequences CUC and, to a lesser extent, AUC and UUC, provided that the conserved base-pairing (Fig. 1a) was maintained. In the experiments described below, the sequence GUC was used as the sole requirement in the target RNA sequence for design of a corresponding enzyme. Future experiments will determine the full extent of the sequence requirements for efficient RNA cleavage.

(2) The region (B) contains sequences highly conserved in naturally occurring cleavage domains. In Fig. 1a, those nucleotides which are conserved in all known self-cleavage domains are specified, while only the positions and required base-pairing of the other nucleotides are indicated. The lengths of the base-paired stem II and associated loop do not appear to be conserved and the loop may be dispensable, as for ASBV. In this case, an RNA enzyme might be active as two halves, or subunits, each held to the other by formation of an extended base-paired stem II. But, conserved tertiary folding within this domain might not be evident from sequence comparisons, being maintained by compensating changes in nucleotide sequence, and initially we have used sequences similar to those of sTobRV for this region.

(3) The catalytic region and RNA substrate are held together by flanking regions of base-pairing (C) which must allow accurate positioning of the enzyme relative to the potential cleavage site in the substrate. The extent and type of base-pairing will directly affect the specificity, affinity (K_m) and turnover (k_{cat}) of an RNA enzyme. For the enzymes described below, a size of eight base-pairs has been arbitrarily chosen for each of the flanking regions (C).

Testing of synthetic ribozymes

To test these principles an indicator gene, chloramphenicol acetyl transferase (CAT)¹⁷, was used to provide RNA transcripts as a target for cleavage by three newly designed RNA enzymes.

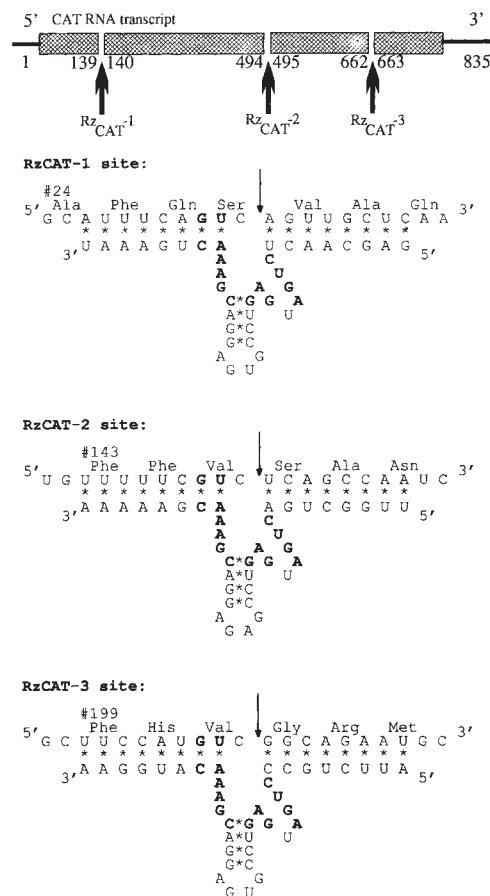


Fig. 4 Design of ribozymes targeted against the CAT gene transcript. Following the design rules described in this paper, ribozymes, termed RzCAT-1, 2 and 3, were targeted against three sites within an 835 base *in vitro* transcript of the CAT gene. The relative locations of the cleavage sites on the transcript are shown schematically with the flanking bases numbered. The three ribozyme sequences are shown with the actual target sequences. Amino-acid sequences of the CAT gene are numbered and the predicted sites for RNA cleavage arrowed. RzCAT-1 and 3 contain 24 base sequences derived from (+) strand sTobRV (region B, Fig. 3), while RzCAT-2 contains a single U to A change in this region.

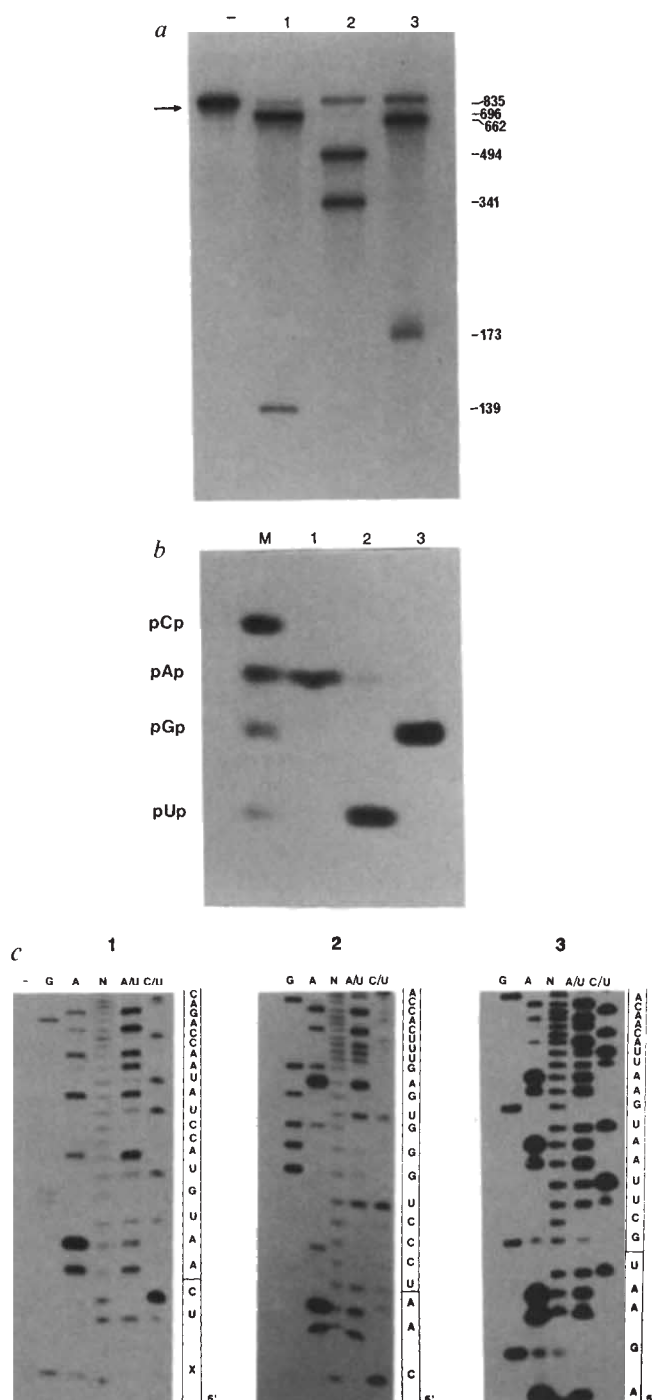
Expression of the CAT gene can provide antibiotic resistance in bacteria, plants and animals and this can be easily assayed. The coding region of the CAT gene transcript was scanned for the occurrence of the sequence GUC as potential sites for RNA cleavage. Three such sites were arbitrarily chosen, and their flanking sequences used to define the eight-nucleotide regions of complementary sequence within the ribozymes, which flanked conserved sequences based on those of the sTobRV (+) strand self-cleavage site (Fig. 4).

The ribozyme sequences (RzCAT 1, 2 and 3) were synthesized as single-stranded oligodeoxyribonucleotides, with the addition of sequences containing half-sites for *EcoRI* (AATTC) and *PstI* (CTGCA) restriction endonucleases at the 5' and 3' termini, respectively. These additional sequences are partly complementary to the overhanging termini left by *EcoRI* and *PstI* digestion of double-stranded DNA. The kinased oligomers were incubated with *EcoRI*-*PstI* digested, phosphatase-treated plasmid DNA in the presence of DNA ligase and DNA polymerase, and after transformation the resulting clones were screened for presence of the *EcoRI*-*PstI* insert. This procedure allowed efficient cloning of the ribozyme sequences using single-strand oligomers, obviating the need for synthesis of a second complementary oligomer.

The ribozyme sequences and the CAT gene were cloned

Fig. 5 Cleavage of the CAT gene transcript. *a*, The [32 P]-labelled CAT RNAs were gel fractionated after incubation alone (–) or with one of the three ribozymes R_{ZCAT}1, 2 and 3 (lanes 1, 2 and 3, respectively). The location of the full-length transcript is arrowed. CAT mRNAs incubated with each of the ribozymes underwent efficient cleavage. In each case only two fragments were produced after incubation with a given ribozyme, and their nucleotide sizes were consistent with the predicted sites for cleavage (that is, 139 and 696, 494 and 341, 662 and 173 base fragments were the 5' and 3' products from R_{ZCAT}1, 2, 3-catalysed cleavage respectively). *b*, 5' Terminal base analysis. The 3' fragments produced by ribozyme cleavage of CAT mRNA were [$5'$ - 32 P]-kinased, gel purified, subjected to complete nuclease digestion and the released terminal residues were fractionated by pH 3.5 PAGE. The 5' terminal nucleotides, determined by reference to markers (lane M), were A, U and G for the fragments produced by R_{ZCAT}1, 2 and 3 (lanes 1, 2 and 3, respectively). *c*, 5' terminal sequence analysis. [$5'$ - 32 P]-labelled CAT mRNA fragments were subjected to base-specific partial ribonucleolytic cleavage. The products were polyacrylamide gel fractionated to allow 5' terminal sequence determination. Panels 1, 2 and 3 correspond to products of R_{ZCAT}1, 2 and 3, respectively. Base specific cleavages are shown above each lane except for lane N which is a size ladder control. Sequences from 4 to 6 to about 25 nucleotides 3' of the cleavage site are indicated for each fragment, and are consistent with the expected sites for cleavage by each ribozyme.

Methods. The CAT gene was obtained from pCM4²⁴ and subcloned as a *Bam*HI fragment into pGEM-blue. This plasmid was linearized with *Hind*III and CAT gene transcripts were obtained using T7 RNA polymerase with 200 μ M [α - 32 P]UTP. Ribozyme sequences were synthesized as oligodeoxynucleotides using automated phosphoramidite chemistry. The 47-, 48- and 44-mer oligonucleotides, R_{ZCAT}1, 2 and 3, respectively, were kinased, ligated with phosphatase treated, *Eco*RI-*Pst*I cut pGEM4 and incubated with the Klenow fragment of DNA polymerase 1 before bacterial transformation. *Eco*RI-linearized plasmids were transcribed with T7 RNA polymerase to produce ribozyme RNAs. Ribozymes were incubated with CAT transcript in 50 mM Tris-HCl pH 8.0, 20 mM MgCl₂ at 50 °C for 60 min, and the products fractionated by 5% polyacrylamide 7 M urea, 25% formamide gel electrophoresis before autoradiography. Ribozyme-cleaved CAT mRNA fragments (2 μ g) were incubated in 50 mM Tris-HCl pH 9.0, 10 mM dithiothreitol with 50 μ Ci [γ - 32 P]ATP and 5 units T4 polynucleotide kinase for 30 min at 37 °C. Radiolabelled fragments were purified on a 5% polyacrylamide gel and digested with an equal volume of 500 units ml⁻¹ RNase T1, 25 units ml⁻¹ RNase T2 and 0.125 mg ml⁻¹ RNase A in 50 mM ammonium acetate pH 4.5 for 120 min at 37 °C. Products were fractionated on a 20% polyacrylamide gel containing 25 mM sodium citrate pH 3.5 and 7 M urea²⁵, and were detected by autoradiography. Terminal sequences were determined directly using the partial enzymatic digestion technique²⁶.



downstream of promoters for T7 RNA polymerase, and this enzyme was used to obtain RNA transcripts *in vitro*. When the 835-nucleotide CAT transcript was incubated with any one of the three ribozymes, efficient and highly sequence-specific cleavage occurred producing two RNA fragments (Fig. 5a). The conditions required for ribozyme-catalysed cleavage were similar to those observed for the naturally-occurring cleavage reactions^{3,5,9,11} with more efficient cleavage occurring at elevated pH, temperature and divalent cation concentrations (data not shown). When present in molar excess, the three ribozymes catalysed almost complete cleavage of the CAT RNA substrate after 60 minutes in 50 mM Tris-HCl pH 8.0, 20 mM MgCl₂ at 50 °C (Fig. 5a). Under similar conditions with 0.1 μ M substrate and 3 μ M ribozyme, the $T_{1/2}$ of CAT mRNA substrate was 3.5, 3.5 and 2.5 min, in the presence of R_{ZCAT} 1, 2 and 3, respectively.

The ribozyme sequences were inactive in the form of oligodeoxyribonucleotides or against the complement of the substrate RNA (data not shown).

The sizes of the CAT RNA fragments produced by each ribozyme were consistent with the predicted sites for cleavage. The 3' terminal cleavage fragments from each ribozyme-catalysed reaction were isolated and 5'- 32 P-kinased. Efficient kinasing of the fragments indicated that they possessed 5' terminal hydroxyl groups, similar to those produced in naturally-occurring cleavage reactions. Terminal nucleotide analysis of these fragments demonstrated that cleavage of CAT sequences by R_{ZCAT} 1, 2 and 3 occurred precisely before nucleotides A, U and G, respectively (Fig. 5b) and RNA sequence determination confirmed that cleavage had occurred at the expected locations within the CAT RNA (Fig. 5c).

Enzymatic catalysis

To demonstrate that these ribozymes cause cleavage of the CAT mRNA substrate in a catalytic manner, each was incubated with a molar excess of substrate, under conditions which should favour both efficient cleavage and product dissociation. Figure 6 shows the results of an experiment where, after 75 min at 50 °C, pH 8.0 in 20 mM MgCl₂, 10 pmol of Rz_{CAT}1 had catalysed specific cleavage of 163 pmol of a truncated CAT mRNA substrate. On average, each ribozyme had participated in greater than 10 cleavage events. Similar results were obtained for Rz_{CAT} 2 and 3 (data not shown), and thus each acts as an RNA enzyme.

Mechanism of cleavage

All three RNA enzymes which were designed according to the principles outlined above have been demonstrated to catalyse effective and precise cleavage of their target RNA sequences. It is therefore likely that these design principles are generally valid, as is the model upon which they are based. The model contains two main assumptions: (1) that catalysis requires that the RNA enzyme and substrate interact through base-pairing (although tertiary interactions presumably exist between enzyme and substrate, they do not need to be considered in the design), and (2) that base-pairing allows precise positioning of a conserved domain (B, Fig. 3) adjacent to a site for cleavage. It may not only be useful, but of functional significance, to consider this conserved domain as the catalytic section or active site of the RNA enzymes.

The RNA enzyme- and related self-catalysed cleavage reactions appear to proceed via concomitant cleavage of the susceptible 3',5' phosphodiester linkage and formation of termini containing 2',3' cyclic phosphodiester and 5' hydroxyl groups^{3,5,9,11}, a mechanism similar to that occurring during base-catalysed hydrolysis of RNA. There exists a previously described case of similar, highly specific RNA cleavage for which detailed structural information is available. During X-ray crystallographic studies of yeast tRNA^{phe}, heavy-atom derivatives were sought. When the tRNA^{phe} crystals or solutions were incubated with lead (Pb(II)) salts, specific cleavage of the RNA occurred between nucleotides 17 and 18 (refs 18, 19). The cleavage was pH dependent, with increased rates observed around neutral pH, and produced termini containing 2',3' cyclic phosphodiester and 5' hydroxyl groups. Three Pb(II) ions were located within the tRNA^{phe} molecule and one, covalently bound to bases 59 and 60 within the TΨC-loop, was positioned in close proximity to the cleaved bond in the D-loop. Brown *et al.*¹⁸ suggested that this Pb(II) ion acted as a source of metal-bound hydroxyl ions near neutral pH, as Pb(II) bound water molecules possess a *pK_a* of ~pH 7.4. The perhaps fortuitously but precisely positioned (Pb-OH)⁺ moiety could abstract a proton from the 2' hydroxyl group of nucleotide 17, thus initiating nucleophilic attack of the phosphate atom in the adjacent phosphodiester bond between nucleotides 17 and 18, and resulting in formation of a 2',3' cyclic phosphate group and 5' hydroxyl leaving group. A similar mechanism has been proposed for the initial step in cleavage by the protein RNase A, where the precisely positioned imidazole of histidine-12 acts as a general base to deprotonate the 2' hydroxyl of the RNA substrate²⁰.

The ribozyme- and self-catalysed cleavage reactions discussed in this paper require divalent metal ions for activity, are pH-dependent, produce similar terminal groups, and thus may involve similar mechanisms to those responsible for (PbII)-catalysed tRNA^{phe} cleavage. In addition, the involvement of such mechanisms would be consistent with the simple structural properties of the ribozymes outlined above. For example, the domain of a highly conserved sequence within the ribozyme might constitute a metal-ion binding site. The flanking base-paired regions joining the ribozyme and substrate would serve to both stabilize the structure of this binding site and bring a bound metal ion in close proximity to the 2' hydroxyl group

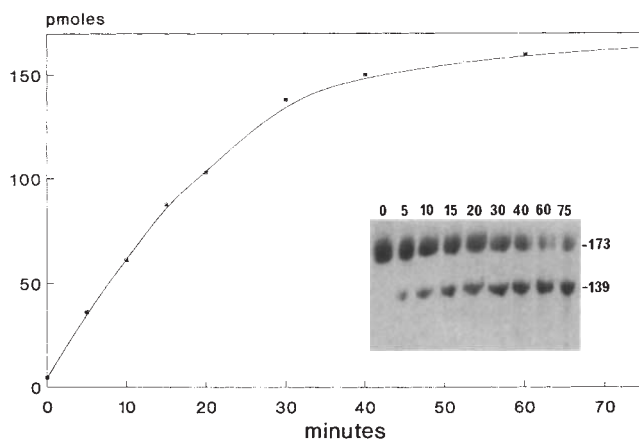


Fig. 6 Enzymatic catalysis. Rz_{CAT}1 (15 pmol) was incubated for varying times with a 20-fold molar excess of truncated (173 base) CAT mRNA containing the site for Rz_{CAT}1 cleavage. Rz_{CAT}1 catalysed cleavage of the CAT RNA to form 5' and 3' fragments, of 139 and 34 bases, respectively. The inset shows the accumulation of the 139 base fragment with time, after PAGE. The amounts of the 139 base fragment were quantified and plotted. After 75 min at 50 °C some non-specific cleavage of RNA was noticed due to the extreme conditions, but 70% of the remaining intact RNAs (163 pmol) had accumulated as the 139 base fragment.

Methods. *Pvu*II-linearized pGEM-blue containing the CAT gene sequence, and *Eco*RI-linearized pGEM4 containing Rz_{CAT}1 were transcribed *in vitro* with T7 RNA polymerase and 200 μM [α -³²P]UTP. 300 pmol of 173 base long, truncated CAT mRNA and 15 pmol of Rz_{CAT}1 were mixed in 100 μl of 50 mM Tris-HCl pH 8.0, 20 mM MgCl₂. 10 μl aliquots were incubated for varying times at 50 °C. Samples were fractionated by 7 M urea, 5% PAGE, stained with toluidine blue, the bands were excised, and amounts of radiolabelled RNA were determined.

adjacent to the susceptible phosphodiester bond. The precisely positioned metal-bound hydroxyl group could then abstract the proton from the 2' hydroxyl group and initiate cleavage of the substrate RNA in a fashion similar to that seen in tRNA^{phe} cleavage. Alternatively, cations may play a structural role in catalysis, stabilizing the interaction between substrate and reactive group(s) present in the RNA enzyme. Detailed structural and kinetic studies are required to test such models for ribozyme action, and may provide a further basis for manipulation of the substrate specificity, and catalytic activity of ribozymes.

Potential applications

The RNA enzymes described in this paper have been shown to be capable of efficient and specific cleavage of RNA sequences *in vitro*, and are based on RNA self-cleavage reactions which occur normally *in vivo*. Accordingly, the ability to design and use specific endoribonucleases for cleavage of particular RNA species or target sequences may prove useful both *in vitro* and *in vivo*.

Ribozymes could be used for the *in vitro* manipulation of RNAs either to produce large quantities of particular RNA fragments or, by using ribozymes with lesser specificities, as a means of physically mapping RNAs. In addition, it may be possible to construct ribozymes as two halves, with the halves held together by extended base-pairing. Each half of such a ribozyme would be specific for one side of a particular cleavage site. By combining ribozyme halves it would be possible to generate new substrate specificities. Given that certain self-catalysed cleavage reactions are reversible to some extent^{3,21}, it may also be possible to isolate and splice together specific RNA fragments using this approach.

A major potential application for these highly sequence-specific endoribonucleases is in cleavage, and thereby inactiva-

tion, of gene transcripts *in vivo*. It is now possible to express foreign genes in a number of bacterial, fungal, plant and animal species either through stable chromosomal transformation of a particular genotype, transient expression via episomal or viral vectors, or by microinjection. Such methods could also be used for the introduction and expression of ribozyme sequences. Genes are universally expressed via RNA transcripts, giving rise to structural/functional RNAs or messenger RNAs for polypeptide synthesized. Essentially any RNA is a potential substrate for cleavage by a ribozyme. Provided that the transcribed sequences of the gene are known, it should be possible to target

one or more ribozymes against specific RNA transcripts. Expression *in vivo* of such ribozymes and cleavage of the transcripts would in effect inhibit expression of the corresponding gene. This 'anti-gene' activity of the ribozymes could provide a basis for various gene and viral therapies and analyses.

We are grateful to Tim Close for pCM4 plasmid and Kim Newell and Paul Whitfeld for synthesis of oligonucleotides. We thank George Bruening for helpful discussion, and Mark Young, Jim Tokuhisa, John Watson, Bob Symons and Jim Peacock for comments on the manuscript. J.H. was supported by a Queen Elizabeth II Research Fellowship.

Received 13 June; accepted 30 June 1988.

1. Cech, T. R. *Science* **236**, 1532-1539 (1987).
2. Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. *Nucleic Acids Res.* **14**, 3627-3640 (1986).
3. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. *Science* **231**, 1577-1580 (1986).
4. Buzayan, J. M., Gerlach, W. L. & Bruening, G. B. *Proc. natn Acad. Sci. U.S.A.* **83**, 8859-8862 (1986).
5. Forster, A. C. & Symons, R. H. *Cell* **49**, 211-220 (1987).
6. Kiefer, M. C., Daubert, S. D., Schneider, I. R. & Bruening, G. B. *Virology* **121**, 262-273 (1982).
7. Bruening, G. B., Gould, A. R., Murphy, P. J. & Symons, R. H. *FEBS Lett.* **148**, 71-78 (1982).
8. Hutchins, C. J. *et al. Plant molec. Biol.* **4**, 293-304 (1985).
9. Epstein, L. M. & Gall, J. G. *Cell* **48**, 535-543 (1987).
10. Buzayan, J. M., Gerlach, W. L., Bruening, G. B., Keese, P. & Gould, A. R. *Virology* **151**, 186-199 (1986).

11. Uhlenbeck, O. C. *Nature* **328**, 596-600 (1987).
12. Forster, A. C. & Symons, R. H. *Cell* **50**, 9-16 (1987).
13. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849-957 (1983).
14. Zaug, A. J., Been, M. D. & Cech, T. *Nature* **324**, 429-433 (1986).
15. Kim, S.-H. & Cech, T. R. *Proc. natn Acad. Sci. U.S.A.* **84**, 8788-8792 (1987).
16. Koizumi, M., Iwai, S. & Ohtsuka, E. *FEBS Lett.* **228**, 228-230 (1988).
17. Alton, N. K. & Vapnek, D. *Nature* **282**, 864-869 (1979).
18. Brown, R. S., Hingerty, B. E., Dewan, J. C. & Klug, A. *Nature* **303**, 543-546 (1983).
19. Brown, R. S., Dewan, J. C. & Klug, A. *Biochemistry* **24**, 4785-4801 (1985).
20. Wlodawer, A., Miller, M. & Sjolín, L. *Proc. natn Acad. Sci. U.S.A.* **80**, 3628-3631 (1983).
21. Buzayan, J. M., Gerlach, W. L. & Bruening, G. *Nature* **323**, 349-352 (1986).
22. Haseloff, J. & Symons, R. H. *Nucleic Acids Res.* **10**, 3681-3691 (1982).
23. Haseloff, J. thesis, Univ. Adelaide (1983).
24. Close, T. J. & Rodriguez, R. *Gene* **20**, 305-316 (1982).
25. Dewachter, R. & Fiers, W. *Analyt. Biochem.* **49**, 184-197 (1974).
26. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1980).

LETTERS TO NATURE

A blue, polarized continuum source near radio galaxy PKS2152-69

S. di Serego Alighieri*, L. Binette†, T. J.-L. Courvoisier*, R. A. E. Fosbury* & C. N. Tadhunter*

* Space Telescope European Coordinating Facility, European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-8046, Garching bei München, FRG

† Canadian Institute for Theoretical Astrophysics, McLennan Physical Laboratories, University of Toronto, 60 George Street, Toronto, Ontario M5S 1A1, Canada

The frequent occurrence of extranuclear emission line regions in radio galaxies and quasars is well known^{1,2}, but the origin and excitation of gas at temperatures of $(1-2) \times 10^4$ K, extending to distances beyond 100 kpc from the nucleus, are not understood in detail. In some cases there is evidence that highly excited gas is associated with the transport of energy from the active nucleus of the galaxy out to the extended radio structures, and in one such case, the radio galaxy PKS2152-69, Tadhunter *et al.*³ reported a blue continuum source coincident with the highest ionization region of a gas cloud, at a projected distance of 8 kpc from the nucleus and lying approximately along the radio axis. Here we report optical observations of the continuum showing that its spectrum rises into the near ultraviolet with $f_\nu \propto \nu^{3.1}$ and that its degree of polarization is about 10% with the electric vector perpendicular to the radius vector from the nucleus. In the absence of a workable intrinsic emission mechanism, we propose that a beamed nuclear source could excite the highly ionized extranuclear cloud and scatter off associated dust to produce the observed continuum.

The radio galaxy PKS2152-69 is characterized by $z = 0.028$, $P_{\text{SGHz}} = 5 \times 10^{25} \text{ W Hz}^{-1}$ with $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The continuum fluxes were measured on images taken with the CCD camera at the 2.2-m ESO/Max-Planck-Gesellschaft telescope on La Silla, Chile, through medium-broad band filters avoiding the brightest emission lines. Table 1 gives details of the observations, which were performed in photometric conditions with

seeing between 1.2 and 2.0 arcsec. Data reduction was done with the ESO MIDAS system⁴. Images taken through the same filter were averaged to eliminate the cosmic ray events, by rejecting pixels deviating by more than five times the expected r.m.s. noise from the mean; the CCD dark current was subtracted; nonuniformities in the response were corrected by means of flat-fields; the sky background was subtracted; and the flux densities were calibrated with the observations of spectrophotometric standard stars⁵. The brightest point of the cloud continuum (marked A in Fig. 1) is closest to the nucleus and has a tail (B) with the same colours extending by 2-3 arcsec in the outward direction. The fainter point source 2 arcsec to the northeast (C) is much redder. Before measuring the cloud flux, the smooth contribution from the outskirts of the galaxy was subtracted by fitting it with a bidimensional polynomial of second degree. In the case of the two ultraviolet filters, the galaxy at the position of the cloud is too faint for a reliable fit and has been subtracted by measuring it through an aperture positioned diametrically opposite to the one used for the cloud.

The continuum flux densities for the cloud listed in Table 1

Table 1 Continuum observations of PKS2152-69

Date	λ_c (Å)	$\Delta\lambda$ (Å)	Exp. (min)	f_λ (cloud) ($10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$)	f_λ (nuc.) ($10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$)
27 Aug. 1986	3,367	169	2×45	16.2 ± 2.0	81 ± 10
29 Aug. 1986	3,471	363	2×45	16.0 ± 1.7	63 ± 8
7 Sep. 1986	4,684	185	3×35	2.1 ± 0.3	195 ± 10
8 Sep. 1986	4,720	183	35	1.9 ± 0.4	185 ± 15
7 Sep. 1986	5,472	272	3×35	1.7 ± 0.3	221 ± 10
8 Sep. 1986	8,190	1,880	3×20	0.3 ± 0.2	201 ± 10

Observations were made with an RCA thinned, backside-illuminated charge-coupled device (CCD) with a scale of 0.35 arcsec per pixel, except for the ultraviolet measurements which used a GEC front-illuminated CCD with a scale of 0.26 arcsec and a fluorescent coating to enhance ultraviolet sensitivity. In the Table, λ_c and $\Delta\lambda$ are the central wavelength and width (full width-half maximum) of the filter. The underlying galaxy and emission line contributions have been subtracted from the flux density of the cloud. The quoted errors include the photon noise of all contributions to the subtracted frame as well as the estimated uncertainty in the photometric calibration.

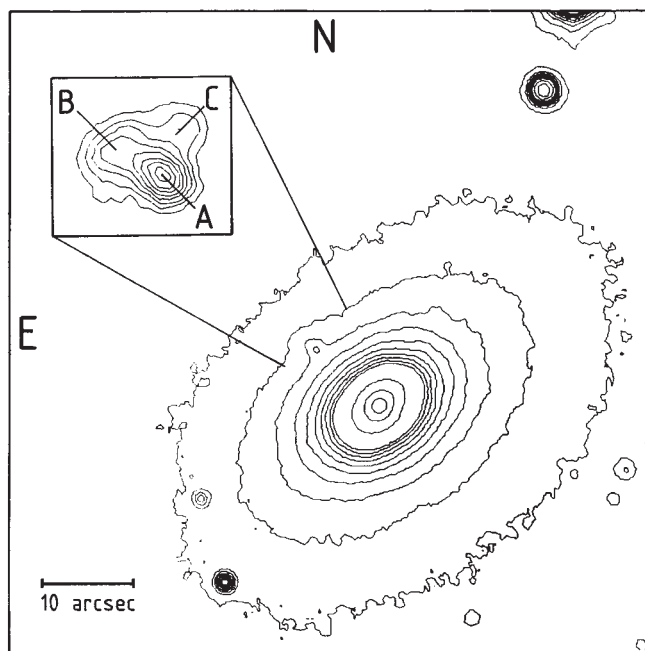


Fig. 1 CCD image of PKS2152-69 through the filter centred at 5,472 Å. The enlargement shows a 3-fold magnification of the cloud region from which the underlying galaxy has been subtracted. North is up and east to the left. The lowest contour level is at $2.35 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1} (\text{arcsec})^{-2}$ ($0.41 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1} (\text{arcsec})^{-2}$ for the enlargement) and the higher ones are spaced by the same amount. To avoid confusion, the four brightest contours at the centre of the galaxy are spaced by a factor of two.

were measured through a circular aperture 7 arcsec in diameter encompassing the whole cloud. The faint red object C, possibly a star, for which we measure a flux density of $0.3 \pm 0.1 \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$ at 5,472 Å, has been subtracted, and the results have also been corrected for the contribution of emission lines within the filter bandpass evaluated using line fluxes measured from calibrated long slit spectra^{3,6}. Emission lines contribute less than 10% of the flux in all filters except for that in the infrared for which, due to the lack of a spectrum in this band, the line contribution was evaluated using a scaled version of the spectrum of NGC3783 (ref. 7), whose line ratios are very similar to those of the cloud in the common range. The continuum flux densities of the cloud presented here agree well with those measured on long slit spectra⁸. Table 1 lists also the flux densities measured through a circular aperture 5.2 arcsec in diameter centred on the nucleus, without correction for emission lines.

Figure 2 shows that the continuum of the cloud is much bluer than that of the nucleus and is continuing to rise in the near ultraviolet. The formal fit of the cloud continuum to a power law ($f_\nu \propto \nu^{-\alpha}$) gives a spectral index $\alpha = -3 \pm 1$. Comparison with continuum free images in the line of [O III] (5,007 Å) and in H α shows that the brightest continuum point is displaced by 1.4 ± 0.2 arcsec towards the nucleus with respect to the peak in the emission lines. Because the tail coincides with this peak, its continuum could contain a contribution due to recombination from the emission line gas. We find that two-photon, $\text{H}^+ \rightarrow \text{H}^0$ and $\text{He}^{2+} \rightarrow \text{He}^+$ recombination accounts for less than 20% of the continuum in the tail. The luminosity of the cloud continuum, integrated over the optical band (3,300–9,000 Å), assuming that it radiates isotropically, is $L_{\text{ctm}} = 6.2 \times 10^{41} \text{ erg s}^{-1}$.

A series of colour maps of the galaxy have been constructed by dividing the calibrated continuum images. These show the presence of a blue, unresolved nucleus which, after removing

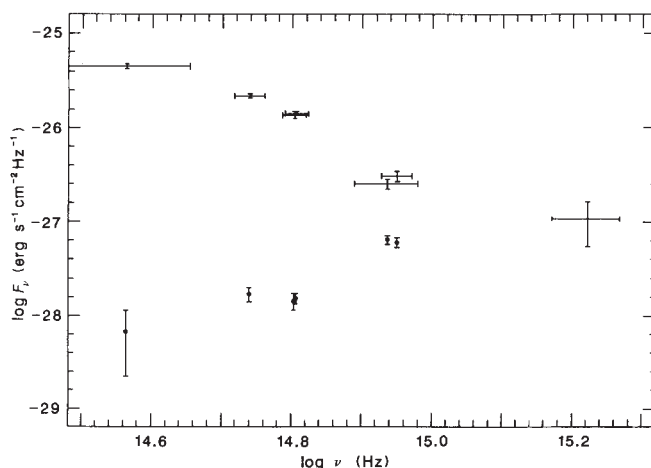


Fig. 2 The continuum flux densities for the nucleus and the cloud (●) of PKS2152-69. The horizontal bars show the width (FWHM) of the filters, while the vertical ones represent the errors. The flux density of the cloud has been corrected for the underlying galaxy contribution and for the emission lines. The rightmost point is a measurement of the nuclear flux taken with IUE⁶.

the underlying stellar light by subtracting a scaled galaxy image in red light from the two shortest wavelength images, has a flux density of approximately $40 \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$. In addition, the southern part of the galaxy is noticeably redder, which may indicate the presence of a band of dust with a major axis in position angle $113^\circ \pm 10^\circ$.

The optical linear polarization of the cloud has been measured in the B-band using observations made on 27 November 1987 at the 3.6 m ESO telescope in La Silla with EFOSC (ESO Faint Object Spectrograph Camera) in its imaging polarimetry mode⁹. For this, a Wollaston prism is inserted into the collimated beam so that each object produces two images perpendicularly polarized and separated by 20 arcsec on the detector (a thinned and backside-illuminated RCA CCD with a scale of 0.675 arcsec per pixel). At the focal plane of the telescope a mask, with alternating opaque and transparent bands 20 arcsec wide, is inserted in order to avoid the overlap of the two beams. Two pairs of 15 min exposures were taken with the instrument rotated by 45° between them.

The data were reduced as described above, but independently for the two beams. The instrumental polarization was corrected using field stars assumed to be unpolarized. The observation of a standard star with known polarization allowed a check on the instrument configuration and on the data reduction procedure, and provided the photometric calibration. After subtraction of the smooth contribution from the outskirts of the galaxy, the degree of linear polarization of the cloud is measured to be $8.6 \pm 1.5\%$ in position angle $124^\circ \pm 5^\circ$. The *E*-vector is then nearly perpendicular to the nucleus-cloud direction, which lies in position angle 42° . These results refer to the whole of the cloud ($\sim 5 \times 5$ arcsec) which, although well resolved, has such a low brightness ($m_B = 20.6$) that we are unable to detect any polarization gradients across it. The brightest point of the cloud in the B-band images, which corresponds to the peak in the continuum images previously described, is definitely polarized as is clearly visible on the raw frames. Since emission lines contribute $27 \pm 3\%$ of the flux in the B band, if these are unpolarized, the linear polarization of the continuum alone is $12 \pm 3\%$.

The detection of an extranuclear source with substantial linear polarization and an energy distribution steeply rising in the near ultraviolet is not readily explained in terms of physical processes usually associated with activity in galaxies. The well known optical jets and radio hot-spot optical counterparts all have

spectra which are steepening in the optical band¹⁰ and so do not appear to extend with any significant power into the ionizing ultraviolet. They do, however, have similar degrees of polarization and are generally interpreted as optical synchrotron radiation.

A prime question to address is that of the relationship between the observed optical continuum and the photon source which is thought to ionize the gas¹¹. The total observed emission line luminosity from the cloud is $L_{\text{lines}} = 2.5 \times 10^{41} \text{ erg s}^{-1}$ (refs 3, 6); 0.5% of this luminosity is in the [Fe X] line, which is coincident with the continuum source. The strongest argument for such an association is the presence of ionization gradients which appear to be centred on this position⁶.

The substantial degree of polarization observed rules out the possibility that all the continuum is produced locally by hot stars. The radial alignment of the geometrical properties of the continuum source in the cloud (its shape, the direction of polarization and the displacement with respect to the emitting gas) suggests that energy is transported from the nucleus to the cloud. Another relevant observation is the wing seen extending some $3,000 \text{ km s}^{-1}$ to the blue of the [O III] emission lines⁶ which may represent momentum transfer to the gas by either a jet or a beam of radiation.

If energy is transported by a light, relativistic jet, we expect a radio hot spot at the position of the cloud. A confirmation of this has to wait for the completion of the Australia Telescope later this year. In this case two emission processes may be envisaged. Synchrotron emission can easily account for the observed polarization, but would imply an unusual electron energy distribution to produce a spectrum rising in the ultraviolet, although the low frequency end of the synchrotron spectrum might be cut off by the Razin effect¹², which is important for $\nu \ll \gamma \nu_{\text{plasma}} \approx 8.9 \times 10^3 \gamma \sqrt{n} \text{ Hz}$ where n is the number of electrons per cubic centimetre and γ is their Lorentz factor. To explain the observed spectrum, $\gamma \sqrt{n}$ must then be at least 10^{11} .

A second possibility is that the observed continuum is the Wien peak produced by inverse Compton scattering of nuclear photons on a nonrelativistic thermal electron population, if the Compton parameter $y = (4kT/mc^2) \max(\tau, \tau^2)$ is much larger than 1. In this case the spectrum peaks at $\nu = 3kT/h$. For a peak in the ultraviolet, as implied by the observations, the electron scattering optical depth τ must be very big ($\tau \approx 100$, for $T = 50,000 \text{ K}$), which makes the polarization difficult to explain.

For either of these possibilities, the production of a blue, intrinsically polarized continuum is a difficulty. The implied luminosity of any source which fits the optical spectrum, especially if it extends to sufficiently high energies to photoionize the cloud, is large. For a black-body fit, the bolometric luminosity is given by $L_{\text{bb}} \approx 4 \times 10^{40} T_4^4 [e^{4.3/T_4} - 1] \text{ erg s}^{-1}$ where T_4 is the black-body temperature in units of 10^4 K . Such a source, if it did ionize the cloud, would produce a higher ionization state than is observed in the low ionization filaments seen extending to similar projected radial distances in the galaxy⁶ if such gas has an electron density which is typical of other well studied extranuclear emission line regions (n_e typically a few hundred cm^{-3} ; ref. 13).

The alternative hypothesis of scattering of collimated optical radiation from the radio galaxy's nucleus does offer an explanation which can explain the observed linear polarization, both in magnitude and direction, and also allows a somewhat redder source spectrum. This model implies that PKS2152-69 is a 'blazar' with a narrow cone of optical radiation intercepting a structure containing both gas and dust. The beam itself would be responsible for ionizing the gas while the dust would scatter a fraction of the radiation in our direction. The choice of an appropriate geometry could explain the polarization with a perpendicular E -vector in a natural way but the scattering may well be inefficient and therefore imply a high beam power, especially if the scattered light is to be substantially bluer than the source spectrum¹⁴ (as would be implied by Rayleigh scatter-

ing from optically thin dust). If we assume that the cloud scatters all the optical radiation from the beam, then the inferred absolute magnitude looking down the beam is only $M_B \approx -23$, which is well within the range for BL Lac objects¹⁵ ($-19 \leq M_B \leq -29$). The tangential ionization gradient⁶ may arise naturally in a beam which has a polar diagram corresponding to synchrotron emission from electrons with a bulk motion Lorentz factor $\gamma \approx 5$. The displacement between the peak in the continuum and that in the lines is more difficult to explain and may imply a peculiar geometrical shape for the dust and gas cloud.

If the scattering hypothesis is correct, it suggests an observational method to detect further members of the class of beamed AGN (blazars) which are not pointing towards us. Searches for blue continuum sources and emission lines along the axis defined by the smallest scale (VLBI) jets should be performed.

This work is based on observations collected at the European Southern Observatory, La Silla, Chile. We thank G. Srinivasan for enlightening discussions. S. di S.A., T.J.-L.C., R.A.E.F. and C.N.T. are affiliated to the Astrophysics Division of the European Space Agency.

Received 27 April; accepted 5 July 1988.

1. Fosbury, R. A. E. in *Structure and Evolution of Active Galactic Nuclei* (eds. Giuricin, G. et al.) 297-308 (Reidel, Dordrecht, 1986).
2. Stockton, A. & McKenty, J. W. *Astrophys. J.* **316**, 584-596 (1987).
3. Tadhunter, C. N., Fosbury, R. A. E., Binette, L., Danziger, I. J. & Robinson, A. *Nature* **325**, 504-507 (1987).
4. *The Midas Users Guide*, Image Processing Group. (European Southern Observatory, Garching, 1987).
5. Oke, J. B. *Astrophys. J. Suppl. Ser.* **27**, 21-35 (1974).
6. Tadhunter, C. N. et al. *Mon. Not. R. astr. Soc.* (in the press).
7. Ward, M. & Morris, S. *Mon. Not. R. astr. Soc.* **207**, 867-881 (1984).
8. Tadhunter, C. N. thesis, Univ. Sussex (1986).
9. Dekker, H. & D'Odorico, S. *ESO Messenger* **46**, 21-22 (1986).
10. Keel, W. C. *Astrophys. J.* **329**, 532-550 (1988).
11. Robinson, A., Binette, L., Fosbury, R. A. E. & Tadhunter, C. N. *Mon. Not. R. astr. Soc.* **227**, 97-114 (1987).
12. Rybicki, G. M. & Lightman, A. P. *Radiative Processes in Astrophysics*, (Wiley, New York, 1979).
13. Heckman, T. M., Baum, S. A., van Breugel, W. & McCarthy, P. J. in *Cooling Flows in Clusters and Galaxies* (ed. Fabian, A. C.) 245-250 (Kluwer, Dordrecht, 1988).
14. White, R. L. *Astrophys. J.* **230**, 116-126 (1979).
15. Véron-Cetty, M.-P. & Véron, P. *A Catalogue of Quasars and Active Nuclei* 3rd edn (ESO, Garching, 1987).

Measurements of atmospheric BrO_x radicals in the tropical and mid-latitude atmosphere

O. N. Singh*, R. Borchers*, P. Fabian*,
S. Lal† & B. H. Subbaraya†

* Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau, FRG

† Physical Research Laboratory, Ahmedabad, India

The possible impact of bromine-bearing components on the Earth's ozone layer has been causing concern¹⁻⁵. Catalytic decomposition of ozone by bromine radicals (BrO_x) is more efficient than by chlorine radicals (ClO_x) and could be enhanced by synergistic effects of BrO_x and ClO_x (refs 6, 7). First vertical profiles of CBrClF_2 and CBrF_3 , two important source gases of atmospheric BrO_x , were measured respectively during 1982-84 and 1980 at 44°N (refs 4, 8). We report profiles measured in 1987 almost simultaneously in the tropics (Hyderabad, 17°N) and at mid-latitudes (Gap, 44°N). Comparison with the earlier data shows that the atmospheric abundances of CBrClF_2 and CBrF_3 , at present $\sim 2 \text{ p.p.t.v.}$ and 1.3 p.p.t.v. , are growing by 12% per year and 5% per year respectively.

‡ Present address: Department of Applied Physics, Institute of Technology, Banares Hindu University, Varanasi, India.

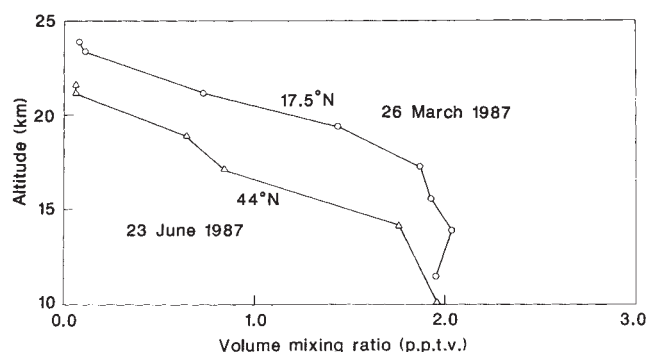


Fig. 1 Vertical profiles of CBrClF₂ measured in 1987 at tropical and mid-northern latitudes.

One of the two significant reaction pathways of BrO + ClO leads to a catalytic cycle for ozone destruction which requires neither sunlight nor atomic oxygen, and the other, although leading to a null cycle, is the only gas phase reaction so far shown to produce OCIO. Only some 20 parts per trillion by volume (p.p.t.v.) of bromine is required to account for the quantities of OCIO measured at McMurdo during austral spring^{7,9}. The above-mentioned reaction is being advanced as a major contributing factor to the Antarctic springtime ozone column loss of ~50% from 1960–86^{7,10–14}. Only few data are available with regard to the atmospheric concentration of the bromine compounds responsible for the bromine budget in the middle atmosphere, and these are preliminary in several cases^{15–18}.

The chief sources of atmospheric bromine are methyl bromide (CH₃Br), of which half is man-made and half is of marine origin, bromoform (CHBr₃), and halons, the chemicals used as non-destructible fire extinguishers. The halons CBrF₃ (CFC-13B1) and CBrClF₂ (CFC-12B1), two of the eight potential ozone depleters, supposedly have the maximum ozone depletion efficiency of 11.4 and 2.7 respectively, which is a few times larger than the relative efficiency of the remaining potential ozone depleters summed together^{19,20}. Like other halons, these are of anthropogenic origin and are used as fire retardants. CFC-13B1 has been rapidly adopted as a fire extinguishant in enclosed spaces such as computer rooms, and CFC-12B1 is increasingly used in portable fire extinguishers. These are mostly banked in fire extinguishing systems which have estimated service lifetimes of up to 40 years. Release into the atmosphere occurs through leakage, fire, inadvertent discharge at disposal, and during test-

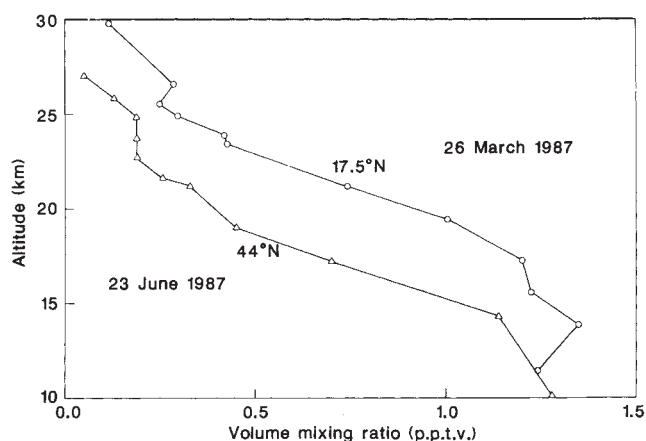


Fig. 2 Vertical profiles of CBrF₃ measured in 1987 at tropical and mid-northern latitudes.

ing and personnel training programs. As fully halogenated hydrocarbons, CFC-13B1 and CFC-12B1 have long lifetimes estimated to range between 62–112 yr (ref. 15) and 29–42 yr (ref. 4) respectively. Their only sink in the atmosphere is ultraviolet photolysis. As the absorption cross-section of CF₃Br is exceedingly small beyond 290 nm ($\sim 10^{-25}$ cm²) (ref. 20), it undergoes barely any photodecomposition in the troposphere and is decomposed mainly in the stratosphere by ultraviolet radiation of 190–285 nm wavelength; CBrClF₂ has relatively larger absorption cross-sections (10^{-22} cm²) beyond 290 nm (ref. 20) and hence its decomposition in the troposphere is not negligible. CFC-12B1 and CFC-13B1 together contribute 10–30% of the organically bound bromine to the atmosphere⁴, augmenting the effect that natural bromine has on ozone depletion.

The tropospheric abundance of CFC-13B1 was first measured by Penkett *et al.*²¹ in 1979 as 0.7 p.p.t.v., and its vertical stratospheric profile was obtained by Fabian *et al.*⁸ in 1980 from air samples collected at 44°N using a cryogenic whole-air sampler; they found stratospheric CBrF₃ mixing ratios decreasing from 1 p.p.t.v. at 10 km to 0.06 p.p.t.v. at 25.9 km. First vertical profiles of CFC-12B1 for the same latitude were obtained by Lal *et al.*⁴, who found CBrClF₂ mixing ratios decreased from 1.5 p.p.t.v. when measured in 1984 at the tropopause to ~0.03 p.p.t.v. at 23 km. They also found, in agreement with surface measurements in the Antarctic²², that tropospheric concentrations increased by 20% per year at that time.

Table 1 Volume mixing ratios VMR in parts per trillion (10^{-12}) of CFCs-13B1 and -12B1 over 17°N and 44°N latitude

CFC-13B1								CFC-12B1			
Hyderabad (26.3.1987)				Gap (23.6.1987)				Hyderabad (26.3.1987)		Gap (23.6.1987)	
Altitude (km)	VMR (p.p.t.v.)			Altitude (km)	VMR (p.p.t.v.)			Altitude (km)	VMR (p.p.t.v.)	Altitude (km)	VMR (p.p.t.v.)
	GC-MS	GC-ECD	mean		GC-MS	GC-ECD	mean				
11.44	1.20	1.27	1.24	10.11	1.36	1.20	1.28	11.44	1.96	10.11	1.96
13.91	1.40	1.29	1.35	14.2	1.21	1.06	1.14	13.91	2.04	14.20	1.76
15.57	1.26	1.18	1.22	17.21	0.80	0.59	0.70	15.57	1.93	17.21	0.84
17.29	1.35	1.05	1.20	18.95	0.41	0.48	0.45	17.29	1.87	18.95	0.64
19.37	0.99	1.02	1.01	21.2	0.41	0.25	0.33	19.37	1.44	21.20	0.06
21.22	0.73	0.75	0.74	21.6	0.27	0.25	0.26	21.22	0.73	21.60	0.06
23.42	0.46	0.39	0.43	22.6		0.19	0.19	23.42	0.110		
23.87	0.50	0.33	0.42	23.75	0.20	0.17	0.19	23.87	0.08		
24.88	0.31	0.28	0.30	24.85		0.19	0.19				
25.55	0.27	0.23	0.25	25.85		0.13	0.13				
26.61	0.36	0.22	0.29	27.1		0.05	0.05				
29.0	0.15	0.08	0.12								

We made two balloon flights in 1987, one over Hyderabad, India (17.5°N) on 26 March, the other on 23 June, over Gap, France (44°N). During each of these flights the MPAE neon-cooled cryogenic sampler^{23,24} was used to collect air samples of 30–50 l volumes at STP at 15 different altitudes between 10–34 km. Along with various other chemical species of interest, the halons CBrClF₂ and CBrF₃ were analysed after recovery of the payload. CBrF₃ was analysed by means of gas chromatography (GC) with an electron capture detector, as well as by Dani 6800 GC/Balzers QMG 511 MS combination, whereas for the analysis of CBrClF₂ GC-electron capture was found to be more sensitive. A 3 m × 2 mm i.d. glass column packed with 10% OV-101 on chromosorb WHP (80–100 mesh) was used for separation for electron capture detection, with temperatures programmed between –60 °C and 70 °C at 10 °C min^{–1} and argon/methane flowing at a rate of 30 ml min^{–1} as the carrier gas. The separation for GC-MS was achieved by a spherosil XOB-075 column operated at temperatures between –50 °C and 150 °C with helium at a flow rate of 30 ml min^{–1} as carrier gas. In both cases a cryotrap preconcentrator was used, consisting of a U-shaped cold trap filled with glass beads and immersed in liquid nitrogen. With sample sizes ranging between 200 and 600 ml, CFC-12B1 and CFC-13B1 could be analysed with a respective precision of ±5% and ±10% or 0.03 p.p.t.v. and 0.10 p.p.t.v., whichever is greater. The absolute calibration was made by means of a two-step static dilution with an accuracy of ±10%. The results are listed in Table 1. The vertical profiles of CFC-12B1 and CFC-13B1 thus obtained are plotted in Figs 1 and 2. For CFC-13B1 (Fig. 2) we have plotted means of the results obtained from both GC-MS and GC-electron capture detector.

By comparing our measurements of the tropospheric concentrations of CFC-12B1 and CFC-13B1 with the values reported earlier, we can obtain the annual growth rates of their tropospheric abundances. But we are reluctant to claim that these are very accurate, primarily because only a couple of measurements are available in each case and furthermore because the concentration of these halons is low with respect to the measuring capability of the analytical equipment used. Even a relatively small uncertainty in any of the above-mentioned measurements may lead to large errors in the growth-rate estimation. When we compare our tropospheric volume mixing ratios of 1.96 p.p.t.v. of F-12B1 with the three values of Lal *et al.*⁴ (September 1982 and 1983, and October 1984) reported earlier, we get the annual growth rates to be respectively ~20%, ~13% and ~12%. In our opinion, the last value seems to be the more realistic. Similarly, from the measurements of Penkett *et al.*²¹ in 1979 and Fabian *et al.*⁸ in 1980 and from our present tropospheric concentration of CBrF₃ of 1.3 p.p.t.v., its annual growth rate is estimated to be little over 4%. This is rather modest compared with the growth rate of CBrClF₂. As CBrF₃ has the longest lifetime of all bromine-bearing species responsible for the bromine budget of the atmosphere, its gradual build up even at smaller emission rates is more probable and lasting

than that of the others. Further, CFC-13B1 has a maximum ozone depletion efficiency more than four times higher than that of F-12B1, which supposedly has the second largest ozone depletion efficiency of all the CFCs (ref. 19).

No information on worldwide release rates of halons into the atmosphere or on global production figures is available. If we take 4.5% as the annual increase of CFC-13B1 in the atmosphere, this would correspond to a global release rate of 1.2 kt per year. The global release rate of CFC-12B1 had been estimated by Lal *et al.*⁴, on the basis of a 20% per year increase in its abundance in the atmosphere, to amount to 3.6 kt per year. If we accept the more likely growth rate of 12% per year, present global emission rates of CFC-12B1 are likely to be about 2.2 kt per year.

In addition to the mid-latitude vertical profiles of CFC-12B1 and CFC-13B1, we have also measured their vertical distributions over the tropics (17.5°N). These could be the first stratospheric profiles of these two important halons determined at tropical latitudes. The volume mixing ratios for CFC-13B1 at lower altitudes are rather similar to those over the two aforesaid latitudes (Fig. 2), but at higher altitudes its concentration in the tropical air is larger and therefore could be measured up to relatively greater heights. Although the tropospheric concentration of F-12B1 over the tropics and at mid-latitudes is exactly the same (Fig. 1), its two vertical stratospheric profiles are very different. In tropical air the concentration of CBrClF₂ decreases gradually with increasing height, but at mid-latitudes there is an abrupt decrease of >50% at ~15 km altitude (Fig. 1). The above-mentioned differences observed between the mid-latitude and tropical profiles are mainly due to tropical upwelling. These two halons, along with all other species originating from tropospheric sources, are carried up to higher altitudes over the tropics through tropical upwelling before they finally decompose and go beyond the detection limit. In most of the profiles we observe a step or wave-like structure. This feature is not new and is being observed in the vertical profiles of almost all the trace gases²⁵.

Although the tropospheric concentration of CFC-13B1, as reported earlier, is considerably less than that of CFC-12B1, CBrF₃ could be measured up to higher altitudes, both over middle and lower latitudes. This is consistent with the absorption cross-sections of these two halons: CBrClF₂ has a larger absorption cross-section by three orders of magnitude compared with CBrF₃ for the ultraviolet radiation beyond 290 nm that is chiefly responsible for the photodecomposition of these substances.

This work was sponsored in part by the German Federal Ministry of Research and Technology. The balloon experiment at Hyderabad was performed under the auspices of ISRO/DFVLR Indo-German collaboration. We particularly thank R. R. Daniel, S. V. Damle and R. U. Kundapurkar of the Tata Institute of Fundamental Research, Bombay, R. T. Redkar of the TIFR Balloon Range at Hyderabad and P. Rajaratnam of the Indian Space Research Organisation (ISRO), Bangalore for their support.

Received 25 April; accepted 11 July 1988.

- Watson, R. T. *The Natural Atmosphere of 1974* (CIAP Monograph 1, Section 5.7.5, Department of Transportation, Washington, DC, 1975).
- Wofsy, S. C., McElroy, M. B. & Yung, Y. L. *Geophys. Res. Lett.* **2**, 215–218 (1975).
- Yung, Y. L., Pinto, J. P., Watson, R. T. & Sander, S. P. *J. Atmos. Sci.* **37**, 339–353 (1980).
- Lal, S., Borchers, R., Fabian, P. & Krüger, B. C. *Nature* **316**, 135–136 (1985).
- Singh, L. N. *Naturwissenschaften* **75**, 191–193 (1988).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1985).
- Pyle, J. A. & Farman, J. C. *Nature* **329**, 103–104 (1987).
- Fabian, P., Borchers, R., Penkett, S. A. & Prosser, N. J. D. *Nature* **294**, 733–735 (1981).
- Tung, K. K., Ko, M. K. W., Rodriguez, J. M., Sze, N. D. *Nature* **322**, 811–813 (1986).
- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Stolarski, R. S. *et al.* *Nature* **322**, 808–811 (1986).
- Solomon, S., Mount, G. H., Sanders, R. W. & Schmeltekopf, A. L. *J. geophys. Res.* **92**, 8329–8338 (1987).

- Hills, A. J., Cicerone, R. J., Calvert, J. G. & Birks, J. W. *Nature* **328**, 405–408 (1987).
- Salawitch, R. J., Wofsy, S. C., McElroy, M. B. *Planet. Space Sci.* **36**, 213–224 (1988).
- Fabian, P., *Handbook of Environmental Chem.* Vol. 4(1) (Ed. Hutzinger, O.) 23–51 (Springer, New York, 1986).
- W.M.O., *Atmospheric Ozone—1985 Global Ozone Research and Monitoring Project*, Report No. 16, Vol. II, 636–37+643 (1985).
- Penkett, S. A., Jones, B. M. R., Rycroft, M. J. & Simmons, D. A. *Nature* **318**, 550–553 (1985).
- Cicerone, R. J., Heidt, L. E. & Pollock, W. H. *J. geophys. Res.* **93**, 3745–3749 (1988).
- Hamitt, J. K. *et al.* *Nature* **330**, 711–716 (1987).
- Molina, M. J., Molina, L. T. & Rowland, F. S. *J. phys. Chem.* **86**, 2672–2676 (1982).
- Penkett, S. A., Prosser, N. J. D., Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **86**, 5172–5178 (1981).
- Rasmussen, R. A. & Khalil, M. A. K. *Geophys. Res. Lett.* **11**, 433–436 (1984).
- Borchers, R. *et al.* *Planet. Space Sci.* **35**, 657–663 (1987).
- Fabian, P. *et al.* *J. geophys. Res.* **84**, 3149–3154 (1979).
- Fabian, P., Borchers, R., Krüger, B. C. & Lal, S. *Geophys. Res. Lett.* **12**, 1–3 (1985).

Structure determination of the new high-temperature superconductor $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$

P. Bordet*, C. Chaillout*, J. Chenavas*, J. L. Hodeau*, M. Marezio*†, J. Karpinski‡ & E. Kaldis‡

* Laboratoire de Cristallographie CNRS, associé à l'Université J. Fourier, 166X 38042 Grenoble Cedex, France

‡ Laboratorium für Festkörperphysik ETH, 8093 Zürich, Switzerland

A new superconducting phase has been identified during measurements of the pressure–temperature–composition phase diagram of the pseudobinary system $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ – O_2 (ref. 1). The new phase has formula unit $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$. Preliminary crystallographic results from electron diffraction showed a doubling of the c -axis¹. Here we report the results of a single-crystal X-ray structural study and measurements of superconductivity in the new compound. The superconducting onset, from a.c. susceptibility measurements, is ~ 40 K. The structure (space group $Ammm$) consists of single $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ -type blocks containing single corner-sharing CuO_4 square-planar chains, alternating with single $\text{YBa}_2\text{Cu}_4\text{O}_8$ -type blocks containing double edge-sharing CuO_4 square-planar chains. This suggests the possibility of similar compounds with other ordered sequences of single and double chains.

The pressure–temperature–composition (P – T – x) phase diagram of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ has been studied for reactive gas (O_2 , N_2 , H_2) pressures of up to 3,000 bar and temperatures of up to 1,700 °C, using a high-purity two-chamber internally heated autoclave². A new phase was identified by its novel X-ray powder pattern at 2,800 bar O_2 and 950–1,200 °C (ref. 1). Single crystals of high perfection and with dimensions of up to 1 mm were grown from high-temperature non-stoichiometric solutions using the travelling-solvent configuration. The use of flux reduces the necessary oxygen pressure dramatically.

The starting material $\text{YBa}_2\text{Cu}_3\text{O}_x$ was analysed volumetrically for oxygen content with an accuracy of ± 0.001 (K. Conder *et al.*, in preparation). Pellets prepared from this material, with oxygen stoichiometry $\text{YBa}_2\text{Cu}_3\text{O}_{6.940}$, and those of a $\text{BaCuO}_2 + \text{CuO}$ mixture of eutectic composition³, were put in an Al_2O_3 crucible in a vertical travelling-solvent configuration. After closing the system and introducing O_2 at a pressure of up to 100 bar, the temperature was increased to 1,060 °C. The system remained under these conditions (manostatic control) for 60 h before cooling to 900 °C at 6 °C h^{−1} and then to room temperature at 60 °C h^{−1}. Single-crystal platelets of the new phase grew within the polycrystalline matrix of the flux and products, and were separated from the flux matrix with difficulty. Details of the phase diagram are given elsewhere (refs 4, 5 and J.K., E.K. and S. Rusiecki, in preparation).

Preliminary a.c. susceptibility measurements performed on single crystals of the new phase reveal a superconducting transition with an onset (T_c) at ~ 40 K. This value seems to be sample dependent: curves exhibiting several drops of susceptibility below ~ 80 K have been obtained for some crystals (Fig. 1). These observations might reflect an inhomogeneity of the oxygen stoichiometry in the samples, or could be due to the intergrowth of the different sequences described below.

Several crystals were taken from a batch and tested with a precession camera equipped with Zr-filtered Mo $K\alpha$ radiation. All crystals investigated had the shape either of a platelet or of a needle, but all exhibited the same diffraction features. The symmetry is orthorhombic, with lattice parameters $a \approx 3.85$ Å, $b \approx 3.87$ Å and $c \approx 50.3$ Å. The systematic absences observed

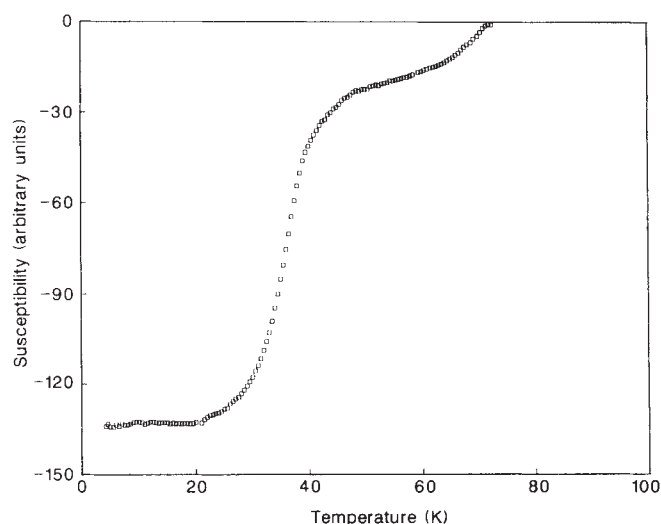


Fig. 1 Plot of a.c. susceptibility of single crystals as a function of temperature. The residual signal at $T < 40$ K is probably due to flux adhering to the crystals.

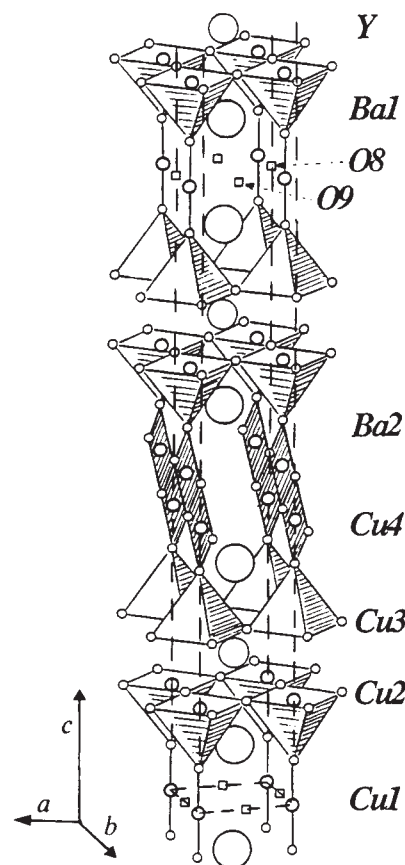


Fig. 2 Schematic drawing of the $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ structure. The partially occupied O8 and O9 sites are represented as squares.

are $k + l - 2n + 1$ for the (hkl) reflections. This implies the space group $Ammm$ or one of its subgroups. The monoclinic phase reported earlier¹ corresponds to the presently described compound, as its cell, when reduced, yields the orthorhombic A -centred unit cell given above.

A trapezoidal platelet was tested with a precession camera and then mounted on a CAD 4 Nonius diffractometer equipped with graphite-monochromated Mo $K\beta$ radiation, which was chosen in order to collect high-angle (001) reflections without

† Also at AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA.

Table 1 Crystallographic data for $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ **(a) Atomic positions, thermal parameters* and occupancy factors**

Atom	Position	x	y	z	U or $U_{11}^\dagger$	U_{22}	U_{33}	n
Ba1	4j	1/2	1/2	0.04310 (3)	0.0141 (5)	0.0147 (5)	0.0094 (4)	1
Ba2	4j	1/2	1/2	0.18797 (2)	0.0059 (4)	0.0063 (4)	0.0070 (4)	1
Y	4j	1/2	1/2	0.11545 (4)	0.0003 (6)	-0.0017 (6)	0.0009 (5)	1
Cu1	2a	0	0	0	0.042 (3)	0.047 (4)	0.020 (2)	1
Cu2	4i	0	0	0.08293 (5)	0.0033 (9)	0.0053 (9)	0.0084 (9)	1
Cu3	4i	0	0	0.14831 (5)	0.0027 (9)	0.0045 (9)	0.0071 (9)	1
Cu4	4i	0	0	0.23012 (5)	0.011 (1)	0.0048 (9)	0.0015 (8)	1
O1	4i	0	0	0.0353 (8)	0.010 (9)	0.09 (3)	0.10 (3)	1
O2	4j	1/2	0	0.0871 (3)		0.5 (2)		1
O3	4i	0	1/2	0.0865 (3)		0.7 (2)		1
O4	4j	1/2	0	0.1430 (3)		0.7 (2)		1
O5	4i	0	1/2	0.1432 (3)		0.8 (2)		1
O6	4i	0	0	0.1937 (2)		0.4 (2)		1
O7	4i	0	1/2	0.2328 (3)		0.8 (2)		1
O8	2b	0	1/2	0		0.6		0.10 (7)
O9	2d	1/2	0	0		0.6		0.20 (7)

(b) Interatomic distances (Å)

Ba1-O1	2.757 (6)	×4	Ba2-O4	2.98 (1)	×2
O2	2.94 (1)	×2	O5	2.96 (1)	×2
O3	2.91 (1)	×2	O6	2.744 (1)	×4
O8	2.899 (2)	×2*	O7	2.96 (1)	×2
O9	2.905 (2)	×2*			
Y-O2	2.402 (9)	×2	Y-O4	2.379 (9)	×2
O3	2.415 (9)	×2	O5	2.378 (9)	×2
Cu1-O1	1.78 (4)	×2	Cu2-O1	2.39 (4)	×1
O8	1.934 (2)	×2‡	O2	1.937 (2)	×2
O9	1.926 (2)	×2‡	O3	1.942 (1)	×2
Cu3-O4	1.944 (2)	×2	Cu4-O6	1.83 (1)	×1
O5	1.951 (2)	×2	O7	1.87 (1)	×1
O6	2.28 (1)	×1	O7	1.939 (1)	×2

* For atoms refined with anisotropic thermal parameters the form of the anisotropic displacement is: $\exp[-2\pi^2(U_{11}h^2a^{*2} \dots + 2(U_{12}hka^*b^* \dots))]$. By symmetry, $U_{ii} = 0$.

† O2-O9 were refined with isotropic thermal parameters (U).

‡ O8 and O9 have occupancy factors of 0.1 and 0.2, respectively.

any overlaps, and to minimize adsorption corrections. The unit-cell parameters determined by refinement of the 2θ positions for 25 high-angle reflections were $a = 3.851(1)$, $b = 3.869(1)$ and $c = 50.29(2)$ Å. All reflections with $k+l=2n$ of the whole reciprocal sphere up to $\theta = 30^\circ$ were measured in the $\bar{\omega}$ -scan mode (1° scan width and 180 s maximum scan time). The total number of reflections was 6,140. Direct methods and Patterson maps were used to find the positions of the Ba, Y and Cu cations; oxygen atom positions were found by Fourier difference syntheses. All positional and isotropic thermal factors were refined on the whole set of data.

An analytical anisotropic absorption correction was applied and the data were averaged in the mmm Laue class. The atomic positions, anisotropic thermal factors of the cations and isotropic thermal factors of the anions were refined simultaneously. As the thermal parameters of Y were too small and those of Cu1 too large, the occupancy factors of these atoms were also refined. As unrealistic results were obtained, perhaps indicative of cation disorder, their values were taken as unity and kept constant together with all other occupancy factors. Difference Fourier syntheses revealed the presence of peaks at the positions $2b$ ($0\frac{1}{2}0$) and $2d$ ($\frac{1}{2}00$). Thus two additional oxygen atoms (O8 and O9) with fixed isotropic thermal factors were placed at these positions and their occupancies were refined. The final refinement cycles, using 761 averaged reflections having intensity $I > 10\sigma(I)$ and unit weights, yielded an R -factor of 5.4% with 47 variables and a chemical formula $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14.30}$. The results, and selected interatomic distances, are listed in Table 1.

Figure 2 shows the structure of $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$: the unit cell contains two formula units. The structure is related to that of

$\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (refs 6-8), and in line with earlier investigations the longest axis of the cell, corresponding to the direction of the cation ordering, is taken as the c -axis. The main difference between the two structures is that in $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ every other layer of corner-sharing CuO_4 square-planar chains is replaced by a layer of double edge-sharing chains running along the b -axis. Similar layers of double CuO_4 chains are found in the structure of SrCuO_2 (ref. 9). Because of the unit-cell A -centring, blocks having the $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ arrangement, separated by blocks containing the SrCuO_2 -type chains, are translated by $b/2$ with respect to each other.

The Cu1 cations at (000) are surrounded by two O1, two O8 and two O9 anions, with the O8 and O9 sites only partially occupied. The occupancy factors of the two sites are 10% and 20% respectively. From the six Cu1-O distances it appears that an octahedral coordination would lead to an effective valence of more than $3+$ for Cu1. By analogy with the $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ ($0 \leq x \leq 1$) structure, it is reasonable to assume that Cu1 is either linearly coordinated by two O1 atoms, or is in a square-planar coordination, with the squares oriented either along the a - or b -axis in a disordered fashion. The large thermal factors might result from the presence of vacancies and/or from displacements of Cu1 from its (000) special position, depending on which of the sites, O8 or O9, is occupied. The Cu1-O1 distance of 1.78 Å is close to that of 1.79 Å found in $\text{YBa}_2\text{Cu}_3\text{O}_6$, which is to be expected in view of the low O8 and O9 occupancies. Note that although along a and b Cu1 has, at least partly, square-planar coordination, the crystal is not twinned because the direction of the b -axis is fixed by the presence of the SrCuO_2 -type chains.

The Cu2 and Cu3 cations have a pyramidal coordination similar to the Cu2 coordination in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$, but in

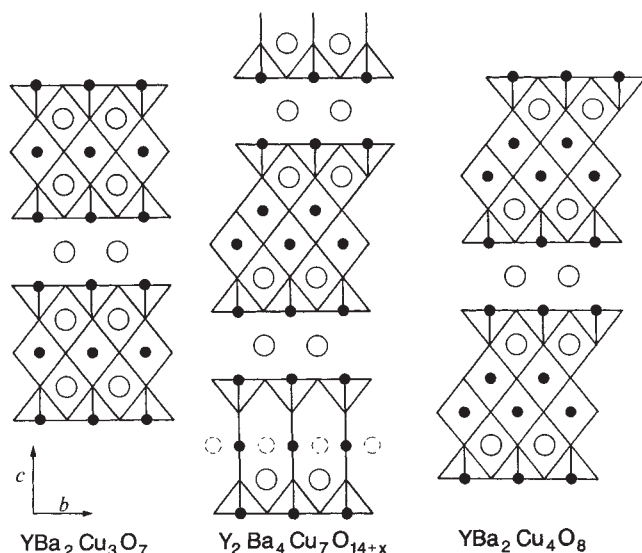


Fig. 3 Schematic representation of the $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$, $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ and $\text{YBa}_2\text{Cu}_4\text{O}_8$ structures. The Cu cations are represented by small full circles, the Y and Ba cations by big open circles and the oxygen anions are located at the line crossings.

$\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ the two CuO_2 square planes that sandwich the Y-layer are no longer equivalent by symmetry. The Cu4 cations are coordinated by one O6 and three O7 anions, forming a slightly distorted square.

Depending on the presence of O8 and O9, the Ba1 cation coordination number can be 8 or 10. This is similar to what happens to the Ba coordination in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ when x changes from 0 to 1. The Y coordination is a slightly distorted cube, with Y–O distances very close to those in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$.

The Y cation coordination is similar to that in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. The small thermal factor observed for this cation is probably due to a Ba substitution on the Y sites. A refinement of the Y occupancy factor yielded a composition of $\text{Y}_{1.7}\text{Ba}_{0.3}$ for this site. The Ba substitution might result from stacking faults in the sample.

Assuming valences of 2+, 3+ and 2– for Ba, Y and O, and the chemical formula $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14.30}$, the average copper valence is 2.09+. The accuracy in the occupancy factors and the cation–anion distances obtained by X-ray diffraction is not sufficient for a precise calculation of the Cu valence from bond length/bond strength rules, but from the relation proposed by Zachariasen¹⁰ one obtains 2.24, 2.24 and 2.51 for the valences of the Cu2, Cu3 and Cu4 cations, respectively. As in $\text{YBa}_2\text{Cu}_3\text{O}_7$, the 3d electrons of copper are not completely localized, although some of the Cu1 cations, being in linear coordination, must have a valence of 1+. The inequivalence of Cu2 and Cu3 mentioned above is not reflected by different values of their valences. A modification of the occupancies of O8 and O9 sites could change the Cu2 valence, while that of Cu3 should be less sensitive to it. The Cu4 valence is rather high, but the Cu valence in SrCuO_2 , where the Cu–O distances are longer, is found to be 2.20+ when calculated by the same method (P. Bendet *et al.*, in preparation).

The new compound $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14.30}$ is intermediate between $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ and $\text{YBa}_2\text{Cu}_4\text{O}_8$, which has been reported to exist only in thin films^{11,12}. The relationship between the three structures is illustrated in Fig. 3. In $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ and in $\text{YBa}_2\text{Cu}_4\text{O}_8$ the chains are either all single or all double, whereas in $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{14+x}$ the single chains alternate with the double chains in an ordered fashion. Compounds may exist which exhibit other ordered sequences of single and double square chains.

We thank Mr H.-P. Staub (ETH) for the susceptibility measurements.

Received 19 July; accepted 28 July 1988.

1. Karpinski, J., Beeli, C., Kaldis, E., Wisard, A. & Jilek, E. *Physica C* **153–155**, 830–831 (1988).
2. Karpinski, J. & Kaldis, E. *J. Cryst. Growth* **73**, 477–483 (1986).
3. Laudise, R. A., Schneemeyer, L. F. & Barnes, R. L., *J. Cryst. Growth* **85**, 569–575 (1987).
4. Karpinski, J. & Kaldis, E. *Nature* **331**, 242–245 (1988).
5. Karpinski, J., Conder, K. & Kaldis, E. *Physica C* **153–155**, 401–402 (1988).
6. Beno, M. A. *et al. Appl. Phys. Lett.* **51**, 57 (1987).
7. David, W. I. F. *et al. Nature* **327**, 310–312 (1987).
8. Capponi, J. J. *et al. Europhys. Lett.* **3**, 1301–1307 (1987).
9. Teske, V. C. L. & Müller-Buschbaum, H. K. *Z. anorg. allg. Chem.* **379**, 234–241 (1970).
10. Zachariasen, W. H. *J. less-common Metals* **62**, 1–7 (1978).
11. Marsh, P. *et al. Nature* **334**, 141–143 (1988).
12. Zandbergen, H. W., Gronsky, R., Wang, K. & Thomas, G. *Nature* **331**, 596–599 (1988).

Periodic area-minimizing surfaces in block copolymers

Edwin L. Thomas, David M. Anderson, Chris S. Henkee & David Hoffman*

Department of Polymer Science and Engineering, and

* Department of Mathematics, University of Massachusetts, Amherst, Massachusetts 01003, USA

An A/B block copolymer consists of two macromolecules bonded together. In forming an equilibrium structure, such a material may separate into distinct phases, creating domains of component A and component B. A dominant factor in the determination of the domain morphology is area-minimization of the intermaterial surface, subject to fixed volume fraction. Surfaces that satisfy this mathematical condition are said to have constant mean curvature (CMC). The geometry of such surfaces strongly influences the physical properties of the material and they have been proposed as candidates for microstructural models in a variety of physical and biological systems. We have discovered domain structures in phase-separated diblock copolymers that closely approximate periodic CMC surfaces. Transmission electron microscopy and computer simulation are used to deduce the three-dimensional microstructure by comparison of tilt series with two-dimensional image projection simulations of 3-D mathematical models. Three structures are discussed here, the first of which is the double diamond microdomain morphology associated to a newly discovered family of triply periodic CMC surfaces¹. Second, a doubly periodic 90° twist boundary between lamellar microdomains, corresponding to a classically known surface (called Scherk's first surface), is described. Finally, we show a lamellar-catenoid microstructure that appears during rearrangement of a lamellar morphology in thin films and is apparently related to a new family of periodic surfaces.

The important energies in the formation of interfaces in block copolymers are the entropic chain-conformational energy and the enthalpic interaction energy, whose strength is characterized by the well-known Flory χ -parameter. For a large value of the product of the polymerization index and χ (the so-called strong-segregation approximation), the interfacial width is small compared with lattice dimensions; minimization of the interfacial contact area becomes the dominant factor. Thus, implicit in the thermodynamics of block copolymer microstructure are conditions that lead to interfacial surfaces of constant mean curvature.

The mean curvature of a surface at a given point is the average of the two principal curvatures and is a measure of the first-order rate of change of area under normal deformation. A surface with nonzero constant mean curvature minimizes area among nearby surfaces produced by local perturbations that preserve the volume (or the volume fraction in the case of periodic surfaces that divide space into two regions) of the regions on either side of the surface. A minimal surface is one whose mean curvature is identically zero; such a surface minimizes area

among all nearby surfaces produced by local perturbations without any volume constraint.

The type of phase-separated domain structure that forms when an amorphous A/B block copolymer is cooled from its high-temperature single-phase state or solution cast from a homogeneous solvent-polymer liquid phase depends on several factors. The most significant is the relative volume fraction of the A and B components. Historically, only three domain types were known for diblock copolymers such as polystyrene-polyisoprene (PS/PI) or polystyrene-polybutadiene (PS/PB) (ref. 2). Spheres make up the minor (A for example) component up to a volume fraction of ~20%. The A component forms cylinders from 20–35%, and alternating A and B lamellae (plane interfaces) from 35–50%, with symmetric results about 50%. Spheres and cylinders are surfaces of nonzero constant mean curvature, whereas planes are minimal surfaces.

New candidates for equilibrium domain structures have recently appeared. Scriven³ was the first to introduce triply periodic minimal surfaces as possible structural models for bicontinuous microstructured fluids. A composite material is said to be bicontinuous if sample-spanning paths exist in any direction in each of the component materials. Subsequent investigations found evidence for triply periodic minimal-surface interfaces in cubic phases of certain amphiphile/water systems (for example, refs 4 and 5). Such an interface defines a bicontinuous structure. This work has paralleled mathematical discoveries of new examples of minimal, and CMC surfaces in recent years (refs 6–8; H. Karcher, N. Kapouleas and J. Pitts, personal communication). Some of these discoveries were motivated by the relevance of these surfaces to microstructural science. Anderson discovered new examples of triply periodic CMC surfaces by numerical simulation. Five families were found, each consisting of deformations of a known triply periodic minimal surface. Three of the families evolve into a limiting configuration that is a close-packing of spheres. H. Karcher and J. C. Nitsche (personal communication) have given mathematical existence proofs for these surfaces.

But although experimental evidence (X-ray scattering) allows determination of the lattice symmetry⁹, it cannot determine if the actual material interfaces are coincident with, or even systematically related to, CMC surfaces with the same periodicity.

Block copolymers provide a favourable opportunity to view directly the partitioning surfaces. In comparison with other microstructured fluids, their length scale is generally larger. Consequently, two-dimensional projection of finite-thickness sections of their structure obtained by transmission electron microscopy (TEM) is less sensitive to overlap problems. Moreover, if one chooses a copolymer, one of whose components has a glass transition temperature well above room temperature (for example, polystyrene has a T_g of 105 °C), the phase-separated structures equilibrated at high temperature can be effectively quenched in and examined at room temperature.

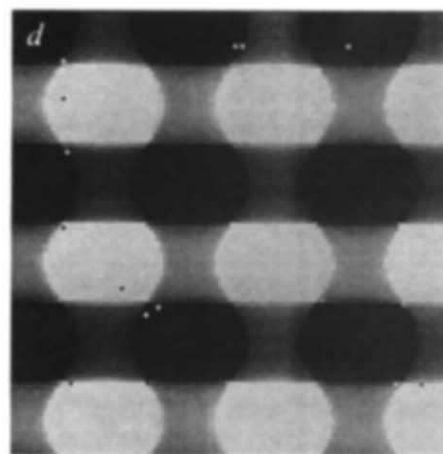
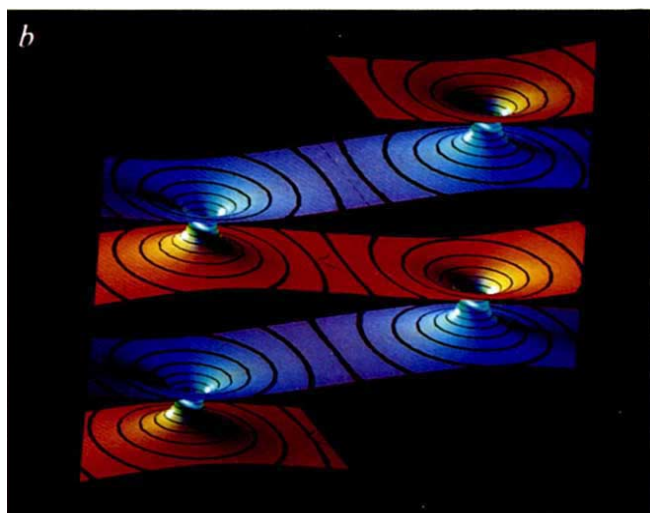
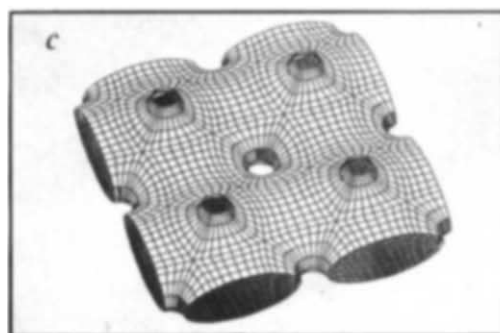
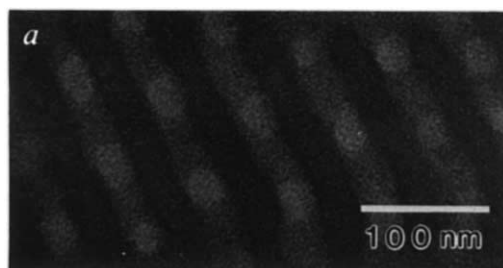
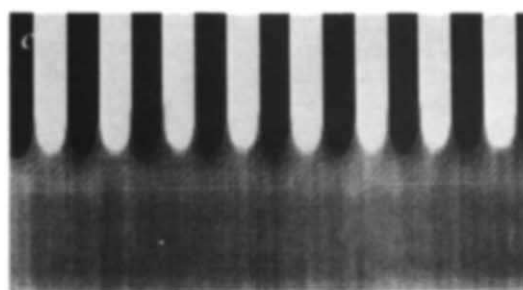
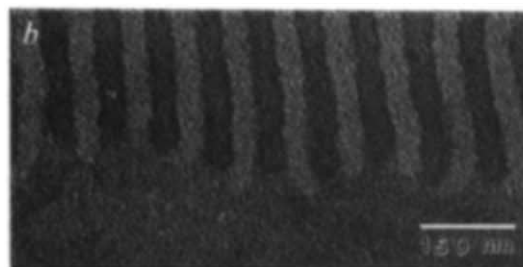
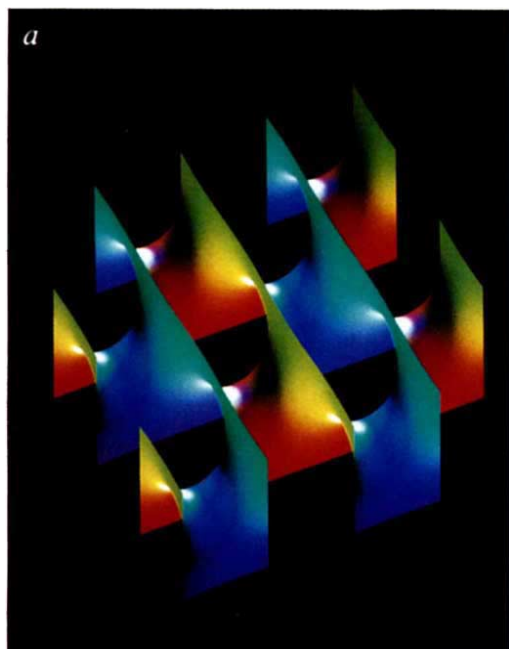
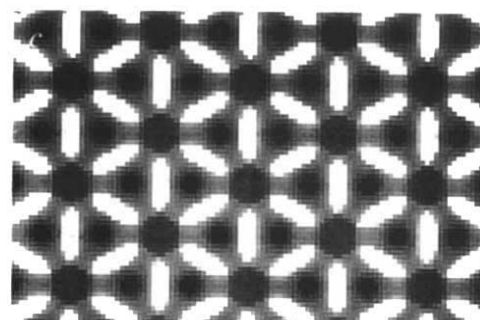
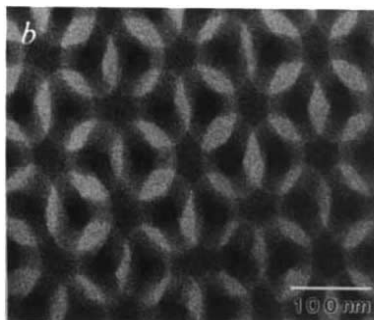
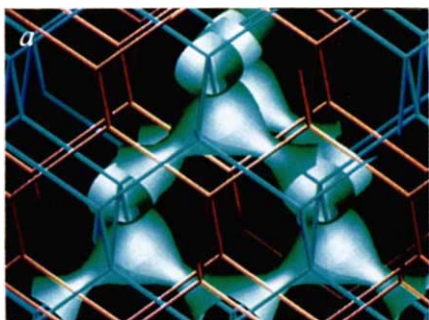
Recently a new microdomain structure, called the ordered bicontinuous double diamond (OBDD) structure was discovered for diblock copolymers of polystyrene-polyisoprene^{10,11}. In the OBDD structure, one phase (for example PI) resides in two intertwined but distinct labyrinthine networks, each exhibiting diamond cubic symmetry, and the other phase (for example polystyrene) resides in the continuous matrix between the two diamond channels. The space group is $Pn3m$ and the matrix is bisected by a connected triply periodic minimal surface (Schwarz's D-surface^{8,12}). The OBDD morphology is observed at volume fractions of 27–38%, corresponding to the region between the cylindrical and lamellar morphologies^{11,13}. The family of CMC surfaces that correspond to this symmetry group^{1,6} extend to a limiting volume fraction of 26.2% for the component in the two diamond channels; furthermore, at this limiting volume fraction, the surface area at fixed lattice parameter exhibits a strong minimum, the dimensionless surface area, $S/V^{2/3}$, for a lattice-fundamental region being lower than

that for any other known triply periodic surface. This is in remarkable agreement with the observation of OBDD at volume fractions near, but always above, 26% in block copolymers.

Electron microscopy and computer simulation can be used to determine the nature of the PS/PI interface^{14,15}. Experimental micrographs can be simulated by sending rays through the three-dimensional theoretical structure (see Fig. 1a) and calculating the resulting two-dimensional projection at each pixel. The projected pixel intensity is determined by testing a large number of evenly spaced points along each ray, finding which material they lie in and then taking the projected intensity at a given pixel to be a linear function of the fraction of points lying within a particular phase (intensities of zero and one for the unstained PS and the OsO₄-stained PI or PB phases respectively). This corresponds to the linear approximation of Beer's law for mass thickness contrast in thin specimens. Figure 1b shows the theoretical [111] projection using the CMC model, and Fig. 1c is the experimental TEM micrograph. The match between the two projections is remarkable. It is not possible to prove mathematically the existence of a specific 3-D surface structure from a finite number of projections, but we have compared a series of experimental projections (images) with computer-simulated projections (tilting from the [111] to the [100] projection). They correspond very closely¹⁰. Other surface structures having the same double diamond symmetry have also been examined¹⁴. For example, a cylindrical-strut double diamond model has been considered. This interface has 6% greater surface area for the same volume fraction. Moreover, simulated projections along the [111] axis for this model give a much poorer fit with the experimental TEM data than does the CMC model, both qualitatively and quantitatively (measured by the mean square error).

Although the detailed arrangement of the microdomains in block copolymer systems has been well studied, the corresponding study of the superstructure of grains of such ordered domains has not. The specific structure of the boundary region between grains is of great importance for physical properties such as modulus and diffusivity that depend on the continuity of the phases across grain boundaries. Thermodynamic considerations, as well as macroscopic properties, strongly suggest that for the lamellar microdomain morphology, phase continuity is maintained across such boundaries¹⁶. For two neighbouring grains consisting of parallel stacks of alternating A and B lamellae, the boundary structure will depend on the orientation of each grain. A situation that is analogous to a 90-degree twist boundary in crystals occurs when the normals to each set of lamellae are orthogonal. This configuration corresponds approximately to a minimal surface discovered in the 1830s by H. Scherk. In essence, the problem is to smoothly join two sets of evenly spaced planes that meet orthogonally. The solution to this problem, called Scherk's First Surface, is shown in three-dimensional perspective, Fig. 2a. The portion of the surface where the two sets of lamellae come into contact consists of a doubly periodic array of saddle-surface regions. Figure 2b depicts a PS/PB block copolymer film containing a region in which the lamellae undergo a twist reorientation of 90 degrees. The characteristic saddle shape of the interface is best observed when viewed along one of the lamellar normals, as is the case for Fig. 2b. Figure 2c shows a correspondingly oriented 2-D simulated projection of Scherk's surface. The detailed agreement strongly suggests that the actual boundary interface between the two grains approximates Scherk's minimal surface, which both affords phase continuity and minimizes unfavourable A/B contact between phases.

Scriven has also suggested periodic area-minimizing surfaces may function as transitory structural pathways for spinodal decomposition and for transitions between equilibrium phase-separated structures¹⁷. Our final example shows regions from a PS/PB block copolymer sample of a partially annealed solution-cast film where the stacked lamellar morphology expected for



◀ **Fig. 1 (Top)** *a*, A computer-generated image of a triply periodic surface of constant mean curvature. This surface belongs to a family of surfaces that includes the classical Schwarz-D minimal surface. The red diamond graph may be surrounded by an identical copy of the surface which surrounds the blue graph. Volume fraction inside each of the diamond channels is 0.131. *b*, A bright-field TEM image of the microdomain morphology of a PS/PB block copolymer. Methods of sample preparation and molecular characterization data are given in ref. 10. *c*, A computer simulated projection of the model structure whose PS/PB interface is Fig. 1*a*. The projection is [111] and the volume fraction of the PS phase (light regions) is 0.62.

Fig. 2 (Middle) *a*, Computer-generated image of Scherk's first minimal surface. It may be considered to be the smooth joining of two orthogonal sets of evenly spaced parallel half-planes. *b*, A bright-field TEM image of the lamellar microdomain morphology of a PS/PB block copolymer in a region where the PS/PB lamellar interface reorients from the plane of the interface to a perpendicular one. Volume fraction of PS is 0.46. *c*, A computer-simulated projection of a smooth structure whose interface is Scherk's First Surface shown in Fig. 2*a*.

Fig. 3 (Bottom) *a*, Bright-field TEM images of a PS/PB block copolymer showing regions exhibiting a modified lamellar microdomain morphology. Volume fraction of PS is 0.45. *b*, Computer generated image of a new minimal surface (consisting of sheet-like regions connected by catenoid-like pieces). *c*, Computer-generated image of a triply periodic minimal surface exhibiting a lamellar-catenoid structure. *d*, A computer-simulated projection of the model structure whose interface is Fig. 3*b*.

this composition has been modified (Fig. 3*a*). Regions for which the lamellar normal lies in the plane of the film and perpendicular to the viewing direction would normally appear as the alternating uniform bright and dark stripes of the lamellar structure. But small regions of alternating layers of PS and PB with regularly spaced catenoidal connectors (hourglass-shaped) are seen, passing from a given phase through the adjacent layer into the same phase to connect each of the phases. This microdomain structure resembles a newly discovered class of periodic minimal surfaces that have sheet-like regions connected by catenoid-like pieces as their component parts (H. Karcher; S. Leibler and T. Maggs, private communications; see Fig. 3*b* and *c*). Figure 3*d* is the simulated image for a projection orthogonal to the layer normal of Karcher's doubly periodic minimal surface.

This research was supported by the National Science Foundation, Divisions of Mathematical Sciences and Materials Research, Polymers, the UMASS-IBM Institute for Interface Science and the Applied Mathematical Sciences Subprogram of the Office of Energy Research, US Department of Energy. The authors thank H. Karcher and S. Leibler for discussions on lamellar-catenoid structures, S. Leibler and T. Maggs for permission to reprint Fig. 3*c* and J. Hoffman for producing Figs 2*a* and 3*b*, *d*.

Received 29 March; accepted 28 June 1988.

1. Anderson, D. M., Scriven, L. E., Davis, H. T. & Nitsche, J. C. *Proc. R. Soc. Phil. Trans.* (in the press).
2. Molau, G. E. *Block Copolymers* (ed. Aggarwal, S. L.) (Plenum, New York, 1970).
3. Scriven, L. E. *Nature* **263**, 123-124 (1976).
4. Longley, W. & McIntosh, T. J. *Nature* **303**, 612-615 (1983).
5. Hyde, S. T., Andersson, S., Ericsson, B. & Larsson, K. Z. *Krist.* **168**, 213-219 (1984).
6. Anderson, D. M. thesis, Univ. Minnesota (1985).
7. Lawson, H. B. Jr *Annls Math.* **92**, 335-374 (1970).
8. Schoen, A. *NASA Tech. Note* TN D-5541 (1970).
9. Charvolin, J., *J. Physique* **46**, 173-183 (1985).
10. Thomas, E. L. et al. *Macromolecules* **19**, 2197-2202 (1986).
11. Hasegawa, H., Tanaka, K., Yamasaki, K. & Hashimoto, T. *Macromolecules* **20**, 1651-1662 (1987).
12. Schwarz, H. A. *Gesamm. Math. Abh.* (Springer, Berlin, 1890).
13. Herman, D. S., Kinning, D. J., Fetters, L. J. & Thomas, E. L. *Macromolecules* **20**, 2940-2942 (1987).
14. Anderson, D. M. & Thomas, E. L. *Macromolecules* (in the press).
15. Callahan, M. J., Hoffman, D. & Hoffman, J. T. *Comm. ACM* **31**, 648-661 (1988).
16. Kinning, D. J., Thomas, E. L. & Ottino, J. M. *Macromolecules* **20**, 1129-1133 (1987).
17. Scriven, L. E. in *Micellization, Solubilization and Microemulsions* Vol. 2 (ed. Mittal, K. L.) (Plenum, New York, 1977).

Glacio-eustatic sea-level control on Red Sea salinity

Robert C. Thunell, Sharon M. Locke & Douglas F. Williams

Department of Geological Sciences, University of South Carolina, Columbia, South Carolina 29208, USA

Previous studies of the Red Sea demonstrated that glacial surface- and bottom-water salinities in the basin were significantly higher than at present. The very low abundance of planktonic foraminifera, the so-called 'aplanktonic zone', during the last glacial indicates that surface-water conditions approached or exceeded the tolerance limits of this plankton group^{1,2}. Glacial sediments are also characterized by high concentrations of magnesium calcite^{3,4}, dolomite³, inorganically precipitated aragonite^{4,5} and benthic foraminifera typical of hypersaline environments^{3,6}. Additional evidence from oxygen isotope records of planktonic^{2,7} and benthic foraminifera^{8,9}, as well as pteropods², demonstrate that glacial-interglacial contrasts in the Red Sea have an amplitude much larger than typically observed in open ocean records. Here we use both oxygen isotope data and a 'frictional overmixing' model to estimate the impact of the most recent (18,000 yr BP) Pleistocene glacio-eustatic sea-level lowering on Red Sea salinity. We estimate that during the last glacial maximum, surface salinities in the central Red Sea were more than 10.0‰ higher than at present. Deep-water salinities were also higher during the last glaciation and remained unusually high through deglaciation. The combination of very high bottom salinities and the onset of pluvial conditions during deglaciation in the Red Sea region prevented ventilation of Red Sea bottom waters and resulted in the accumulation of organic-rich sediments.

The hydrographic character of a marginal or semi-enclosed basin is strongly influenced by the geometry of the strait that connects it to the open ocean¹⁰⁻¹⁴. Significant changes in sea level alter the volume transport through a strait, and in turn modify the water mass properties of the marginal basin as the salt balance between the open ocean and the marginal sea is altered.

Here we estimate quantitatively glacial surface salinities in the Red Sea by two independent approaches. First, we use oxygen isotope records for the planktonic foraminifera *Globigerinoides ruber* (150-300 µm size fraction) from Red Sea Core CH61-153 and Gulf of Aden Core RC9-166 to determine past surface salinities empirically (Table 1). Second, we use the 'frictional overmixing' model of Assaf and Hecht¹¹ to estimate glacial salinities, taking into account the effect of glacial sea-level lowering on the geometry of the Bab-el-Mandeb strait. Like most mixing models that describe water exchange between the open ocean and a marginal sea^{11,13,14}, the model of Assaf and Hecht¹¹ is based on the principle of conservation of mass and salt, and uses such parameters as strait dimensions and net evaporation over the marginal basin. The application of these models to basins such as the Red Sea^{11,13}, the Mediterranean Sea¹⁴ and the Black Sea¹¹ results in predicted salinities consistent with observational data. They thus provide a reasonable means for estimating palaeosalinities.

Red Sea core CH61-153 was collected from 19°43' N, 38°41' E at a depth of 2,704 m, and was not from a hot brine pool¹⁵. The chronology for this core is based primarily on two radiocarbon dates of planktonic foraminifera⁵ and an assumption of constant sedimentation rates between these dates (Table 1). Typically of Red Sea cores, CH61-153 contains an 'aplanktonic zone' during the last glacial and as a result a number of samples from this interval lacked sufficient planktonic foraminifera for isotopic analysis (Fig. 1, Table 1). Gulf of Aden core RC9-166 was recovered from 12°09' N, 44°23' E at a depth of 739 m. No

Table 1 Red Sea and Gulf of Aden data

Ch61-153 (Red Sea)					RC9-166 (Gulf of Aden)		
Core depth (cm)	Est. age (kyr BP)	$\delta^{18}\text{O}$ <i>G. ruber</i> (‰)	Org. carbon (%)	Foram. number ($\times 100 \text{ g}^{-1}$)	Core depth (cm)	Est. age (kyr BP)	$\delta^{18}\text{O}$ <i>G. ruber</i> (‰)
10	0.6	-2.40	0.6	14.3	20	0.9	-2.55
20	1.1	-2.84	0.5	30.0	30	1.9	-2.70
30	1.7	-3.54	0.4	19.1	40	2.9	-2.71
40	2.3	-3.85	0.3	52.5	50	3.8	-3.57
50	2.9	-2.90	0.3	13.9	70	5.7	-3.01
60	3.4	-2.54	0.4	31.9	80	6.7	-3.64
70	4.0	-3.81	0.4	31.8	90	7.6	-2.49
80	4.6	-3.60	0.3	7.3	100	8.6	-2.41
90	5.2*	-2.82	0.3	3.5	110	9.6	-2.17
100	5.7	-2.53	0.3	3.8	120	10.5	-1.65
110	6.2	-2.83	0.4	38.5	130	11.3	-1.85
120	6.8	-1.90	0.4	1.3	140	12.2	-1.39
130	7.3	-3.40	0.4	15.6	150	13.0	-1.07
140	7.8	-3.65	0.3	22.2	160	13.8	-1.04
150	8.4	-2.52	1.2	0.2	170	14.7	-0.93
160	8.9	-3.55	1.3	0.8	180	15.5	-1.15
170	9.5	-2.97	1.5	1.3	190	16.3	-1.10
180	10.0	-3.22	0.4	1.2	200	17.2	-0.96
190	10.9	-0.93	0.2	0.9	210	18.0	-0.62
200	11.7	-0.77	0.2	1.0	220	18.9	-0.88
210	12.6	-0.04	0.2	1.7	230	19.7	-1.49
220	13.4	-0.55	0.4	0.9	240	20.6	-1.08
230	14.3	-0.52	0.2	2.1			
240			0.4	0.9			
250			0.2	2.2			
260	18.1	1.57	0.2	0.7			
270	18.9*	0.67	0.3	3.0			
280	19.7		1.1	10.8			
285	20.1	1.08	0.4	7.7			
290	20.5	0.25	0.3	11.6			

* Radiocarbon ages from ref. 5.

radiocarbon dates are available for this core for absolute age control, and so the stratigraphy is based simply on identifying the 18,000 yr BP level in the oxygen isotope record and assuming a zero age for the core top and a constant sedimentation rate (Fig. 1, Table 1). We realize the limitations of this approach in accurately dating individual events in each of the oxygen isotope records. But for the present study it is critical that we be able to identify the 18,000 yr level and we feel that this has been done satisfactorily.

The *G. ruber* oxygen isotope records show that the differences in $\delta^{18}\text{O}$ between the last glacial maximum and the present in the Red Sea and the Gulf of Aden are approximately 4.0‰ and 2.0‰, respectively (Fig. 1). The glacial/interglacial change in $\delta^{18}\text{O}$ of ocean surface water due to global ice volume changes has been estimated to be between 1.1 and 1.5‰¹⁶⁻¹⁹. Absolute $\delta^{18}\text{O}$ values for the Red Sea and Gulf of Aden are more similar during the latter half of the Holocene than during the last glacial (19,000–14,000 yr BP), when Red Sea $\delta^{18}\text{O}$ values were 0.5–2.0‰ heavier than in the Gulf of Aden (Fig. 1). This large difference in $\delta^{18}\text{O}$ values must reflect increased glacial salinities and/or decreased surface water temperatures in the Red Sea relative to the Gulf of Aden. Likewise, the 2.5–3.0‰ change in the Red Sea oxygen isotope record (Fig. 1) in excess of the 1.1–1.5‰ glacial/interglacial ice-volume effect indicates significant temperature and/or salinity changes in this basin during the past 20,000 yr.

If the 2.5–3.0‰ excess change in $\delta^{18}\text{O}$ were due solely to temperature, it would require an unrealistic decrease of 12–15 °C in Red Sea glacial surface-water temperatures relative to present values²⁰. According to CLIMAP reconstructions^{21,22}, sea surface temperatures in the northern Indian Ocean and Gulf of Aden decreased by ≤ 2 °C during the last glacial maximum.

The 2.0‰ glacial/interglacial $\delta^{18}\text{O}$ change observed in the Gulf of Aden is consistent with a 1.5‰ ice volume effect and a 2 °C decrease in surface waters. Considering the present distribution of sea surface temperatures in this region, it seems unlikely that the central Red Sea would have been 10 °C colder than the Gulf of Aden during the last glacial. At present, sea surface temperatures in the central and southern Red Sea are similar to those in the Gulf of Aden during the summer and warmer in the winter²³.

The difficulty in attributing the excess ^{18}O enrichment in the Red Sea to glacial temperatures suggests that the large-amplitude $\delta^{18}\text{O}$ changes are due to salinity changes. The salinity difference in the Red Sea between now and the last glacial maximum can be calculated by determining the empirical relation between the oxygen isotopic composition of foraminiferal calcite and surface-water salinity in the Red Sea. Craig²⁴ has determined the relation between the isotopic composition of sea water and salinity, through measurements of Red Sea surface and deep waters, as

$$\delta w = 0.29S - 10.08 \quad (1)$$

where δw is the oxygen isotopic composition of sea water (relative to PDB) and S is salinity. The slope of the Red Sea line is about half that of the equation determined for the salinity-oxygen-isotope relation in North Atlantic waters²⁵. The isotopic composition of seawater is involved in the isotopic composition of foraminiferal calcite according to the palaeotemperature equation of Craig²⁰:

$$T = 16.9 - 4.2(\delta c - \delta w) + 0.13(\delta c - \delta w)^2 \quad (2)$$

where T is the temperature of seawater in °C, δw is the $\delta^{18}\text{O}$ of seawater (relative to PDB), and δc is the $\delta^{18}\text{O}$ of calcite

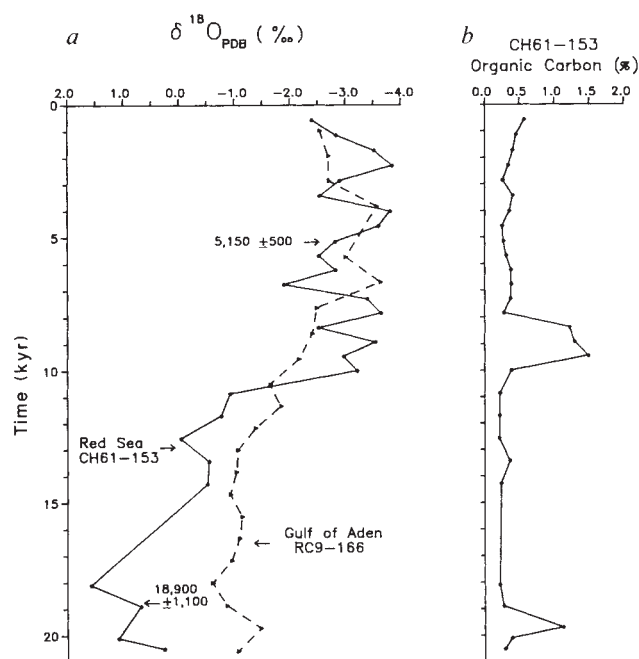


Fig. 1 a, Comparison of the oxygen isotopic records for Red Sea core CH61-153 (solid line) and Gulf of Aden core RC9-166 (broken line). Both records are based on analyses of the planktonic foraminifera *Globigerinoides ruber*. The radiocarbon ages are from ref. 5. b, Organic-carbon content for Red Sea core CH61-153.

(relative to PDB). Substitution of equation (1) into equation (2) and solving for S gives

$$S = 34.76 + \frac{\delta c}{0.29} - \left(\frac{4.2 - \sqrt{8.85 + 0.52T}}{0.08} \right) \quad (3)$$

If certain assumptions are made about surface temperature, equation (3) can then be used to estimate glacial salinities. Two different models are considered in making this calculation. If we assume no surface water temperature difference between glacial and the present in the Red Sea (that is, if all the excess in the glacial/interglacial $\delta^{18}\text{O}$ record is attributed to salinity change), then glacial salinity is estimated to have been 50.1% in the central Red Sea, or 11.1% higher than present. Alternatively, if we assume, based on the reconstructions for the northern Indian Ocean^{21,22}, that surface water temperatures in this region during the last glacial were 2°C colder than at present, the estimated glacial salinity would have been 48.6%. These estimated glacial salinities are slightly higher than the 46.4% previously calculated by Deuser *et al.*⁷ using oxygen isotopic data.

The overmixing model developed by Assaf and Hecht² can be used to examine independently the reasonableness of the above salinity estimates. The present salinity difference between inflowing Gulf of Aden water and outflowing Red Sea water across the Bab-el-Mandeb strait is related to the strait dimensions and net evaporation according to the equation:

$$1 - R = Q^{2/3} L^{1/3} W^{-2/3} D^{-4/3} \quad (4)$$

where R is the ratio of the inflowing (S_i) to outflowing (S_o) salinities, Q is the net evaporation ($\text{m}^3 \text{s}^{-1}$), and L , W and D are the length, width and depth of the strait in metres, respectively. The model is obviously very sensitive to sea level change, as this variable controls the volume transport through the strait. The model predicts that a reduction in width or depth of the strait or an increase in net evaporation (excess over precipitation and runoff) will result in an increase in the salinity difference

between a sill-enclosed basin and the open ocean.

For the last glacial maximum we have used an estimate of ~ 120 m for the sea level lowering^{26,27}, although values ranging from -80 to -150 m have been proposed²⁸. Under these conditions the Bab-el-Mandeb strait dimensions would have been reduced to $L = 1.6 \times 10^5$ m, $W = 1.2 \times 10^4$ m and $D = 60$ m. A 120-m decrease in sea level would have also reduced the surface area of the Red Sea by $\sim 50\%$ (ref. 8). Thus, the net evaporation Q would have been the result of evaporation over a substantially smaller surface area. Assuming an evaporation loss similar to today ($6.34 \times 10^{-8} \text{ m s}^{-1}$)²⁹ and a glacial surface area of $2.19 \times 10^{11} \text{ m}^2$ yields a glacial net evaporation rate of $1.388 \times 10^4 \text{ m}^3 \text{s}^{-1}$. Using these values in equation (4) gives $R = 0.75$. If, as suggested by Duplessy¹⁶, glacial salinities in the Gulf of Aden were similar to present values (36.7‰), the predicted salinity of outflowing glacial Red Sea water would have been $S_o = 36.7/0.75$, or $\sim 48.9\%$. By comparison, glacial sea level lowerings of 80 and 150 m yield salinities of 42.2 and 101.9‰, respectively, for glacial Red Sea outflow water.

The salinity estimate derived from the overmixing model using a glacial sea level of ~ 120 m agrees well with the isotopically determined palaeosalinity estimate of 48.6‰ that assumed a 2°C temperature drop. However, the salinity calculated from the overmixing model is made assuming a glacial evaporation rate similar to that of the present. Pollen records³⁰ and quartz distributions³¹ in the Arabian Sea indicate increased aridity on the Arabian continent during the last glacial. Similarly, low lake levels indicate very arid conditions in Africa at this time³². As a result, the glacial evaporation rate over the Red Sea was probably higher than at present and the glacial salinity calculated from the overmixing model may be slightly underestimated.

What were the consequences of glacial salinities of 48.6–50.1% in the Red Sea? Recent laboratory studies³³ indicate that the upper salinity limits for the planktonic foraminiferan *G. sacculifer* and *G. ruber* are ~ 47 and 49‰, respectively. These are the two most abundant planktonic foraminiferal species in the Red Sea during the late Pleistocene^{1–3}. The 'aplanktonic' interval^{1,2} in the Red Sea, which persisted throughout the last glacial, is most likely an ecological response to greatly elevated surface salinities. Similarly, the presence of dolomite and inorganically precipitated aragonite in glacial age sediments from the Red Sea can be attributed to very high salinities^{3–5}.

In addition to producing the widespread 'aplanktonic' zone, the high salinities in the glacial Red Sea may have 'preconditioned' the basin for subsequent deposition of organic-rich sediments at around 9,000–10,000 yr BP^{34,35} (Fig. 1). As the connection between the Red Sea and the Gulf of Aden was re-established during deglaciation, surface water salinities in the Red Sea would have gradually decreased. However, high concentrations of Mg calcite^{3,4}, dolomite³ and inorganically precipitated aragonite⁴ indicate that very high salinities persisted in Red Sea deep waters until about 10,000 yr BP. Pollen and lake level records for North Africa indicate high rates of rainfall from 12,000 to 8,000 yr BP³², with pollen records from the Arabian coast³⁰ also suggesting very wet conditions at the beginning of the Holocene. This interval of increased precipitation over Africa and Arabia corresponds to a period of intensification of the southwest monsoon^{36–41}. Rossignol-Strick⁴² has recently shown that stream and wadi discharge into the Red Sea was then greatly increased, resulting in a lowering of surface salinities. The rapid and large-scale decrease in Red Sea $\delta^{18}\text{O}$ values ($\sim 2.0\%$) between 11,000 and 10,000 yr BP (Fig. 1) is most probably the result of this increased freshwater input and reduced surface salinities. The combination of very high deep-water salinities and increased precipitation/runoff would have resulted in a stable density stratification of the water column. If the salinity contrast between surface and deep waters was high enough, all the inflowing water from the Gulf of Aden would have left the Red Sea as a subsurface current and very little new bottom water would have been formed⁴³. This circulation regime would

be analogous to the present summer flow patterns between these basins⁴³.

A reduction or elimination of bottom-water formation at around 10,000 yr BP would have resulted in a rapid depletion of bottom-water oxygen. In addition, the increased runoff at this time may have raised surface water nutrient concentrations and increased productivity. The combination of decreased bottom-water ventilation and increased organic carbon flux would explain the accumulation of organic-rich sediments at this time.

We thank Rick Fairbanks, Jan Jenkins and Lloyd Keigwin for their comments on of this manuscript. Core material was supplied by the Woods Hole Oceanographic Institution and Lamont-Doherty Geological Observatory. This research was partly supported by the US NSF.

Received 4 April; accepted 5 July 1988.

1. Berggren, W. A. & Boersma, A. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 282-298 (Springer, Berlin, 1969).
2. Reiss, Z. *et al. Quat. Res.* **14**, 294-308 (1980).
3. Locke, S. thesis, Univ. South Carolina (1986).
4. Milliman, J. D., Ross, D. A. & Ku, I. L. L., *J. sedim. Petrol.* **39**, 724-736 (1969).
5. Ku, T. L., Thurber, D. L. & Mathieu, G. G. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 348-359 (Springer, Berlin, 1969).
6. Winter, A. *et al. Mar. Geol.* **53**, 17-22 (1983).
7. Deuser, W. G., Ross, E. H. & Waterman, L. S. *Science* **191**, 1168-1170 (1976).
8. Luz, B. & Reiss, Z. *Utrecht micropaleont. Bull.* **30**, 129-140 (1983).
9. Almogi-Labin, A., Luz, B. & Duplessy, J. C. *Paleogeogr. Paleoclim. Paleoecol.* **57**, 195-211 (1986).
10. Stommel, H. & Farmer, H. J. *mar Res.* **12**, 13-20 (1953).

11. Assaf, G. & Hecht, A. *Deep Sea Res.* **21**, 947-958 (1974).
12. Bethoux, J. P. *Oceanol. Acta* **2**, 157-163 (1979).
13. Anati, D. A. *Oceanol. Acta* **3**, 395 (1980).
14. Bryden, H. L. & Stommel, H. M. *Oceanol. Acta* **7**, 289-296 (1984).
15. Ross, D. A. & Degens, E. T. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 363-367 (Springer, Berlin 1969).
16. Duplessy, J. C. *Nature* **295**, 494-498 (1982).
17. Shackleton, N. J. *Phil. Trans. R. Soc. B* **289**, 169-182 (1977).
18. Mix, A. C. & Pisias, N. G. *Nature* **331**, 249-251 (1988).
19. Labeyrie, L., Duplessy, J. C. & Blanc, P. *Nature* **327**, 477-482 (1987).
20. Craig, H. & Gordon, L. I. in *Stable Isotopes in Oceanographic Studies and Paleotemperatures* (ed. A. Tongiorgi) 1-122 (Consiglio Nazionale delle Ricerche, Laboratorio di Geologia Nucleare, Pisa, 1965).
21. CLIMAP Project Members *Science* **191**, 1131-1137 (1976).
22. CLIMAP Project Members *Geol. Soc. Am. Map and Chart Series* MC-36 (1981).
23. Siedler, G. *Meteor. Forsch. Ergebn. R. A.* **4**, 1-76 (1968).
24. Craig, H. *Science* **154**, 1544-1548 (1966).
25. Epstein, S. & Mayeda, T. *Geochim. cosmochim. Acta* **4**, 213-284 (1953).
26. Milliman, J. D. & Emery, K. O. *Science* **162**, 1121-1123 (1968).
27. Chappell, J. & Shackleton, N. J. *Nature* **324**, 137-140 (1986).
28. Matthews, R. K. *Dynamic Stratigraphy* (Prentice-Hall, Englewood Cliffs, 1984).
29. Maillard, C. & Soliman, G. *Oceanol. Acta* **9**, 249 (1986).
30. Van Campo, E., Duplessy, J. C. & Rossignol-Strick, M. *Nature* **296**, 56-59 (1982).
31. Kolla, V. & Biscaye, P. E. *J. sedim. Petrol.* **47**, 642-649 (1977).
32. Street, F. A. & Grove, A. T. *Quat. Res.* **12**, 83-118 (1979).
33. Hemleben, C. in *II Int. Conf. Paleoceanogr., Woods Hole* (1986).
34. Herman, Y. in *VIII Congr. Int. Ass. Quat. Res.* 325-348, Salt Lake City (1968).
35. Ivanova, E. V. *Mar. Micropaleontol.* **9**, 335-364 (1985).
36. Ritchie, J. C., Eyles, C. H. & Haynes, C. V. *Nature* **314**, 352-355 (1985).
37. Cullen, J. L. *Paleogeogr. Paleoclimatol. Paleoecol.* **35**, 315-356 (1981).
38. Kutzbach, J. E. *Science* **214**, 59-61 (1981).
39. Prell, W. L. in *Milankovitch and Climate Part I* (eds Berger, A. I. *et al.*) 349-366 (Reidel, Dordrecht, 1984).
40. Rossignol-Strick, M. *Nature* **304**, 46-49 (1983).
41. Rossignol-Strick, M. *Paleogeogr. Paleoclimatol. Paleoecol.* **49**, 237-263 (1985).
42. Rossignol-Strick, M. *Paleoceanography* **2**, 333-360 (1987).
43. Wyrtki, K. in *Colloq. Int. CNRS* no. 215 (*Processus de Formation des Eaux Océaniques Profondes*) 91-106 (1974).

Estimation of heat and chemical fluxes from a seafloor hydrothermal vent field using radon measurements

N. D. Rosenberg*, J. E. Lupton*†, D. Kadko‡, R. Collier‡, M. D. Lilley§ & H. Pak†

* Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA

† Marine Science Institute, University of California, Santa Barbara, California 93106, USA

‡ School of Oceanography, Oregon State University, Corvallis, Oregon, 97331, USA

§ School of Oceanography, University of Washington, Seattle, Washington 98195, USA

The circulation of seawater through newly formed ocean crust at mid-ocean ridge spreading centres is important in the oceanic heat and chemical budgets. The heat transfer in submarine hydrothermal systems accounts for ~25% of the total global heat loss¹. The transfer of mass resulting from basaltic alteration affects the geochemistry of seawater² and is responsible for the formation of ore deposits on and beneath the seafloor³. Various geophysical techniques have been employed in efforts to determine the heat output from hydrothermal vent fields⁴⁻⁹; however, the magnitudes of the heat and chemical fluxes through these systems remain uncertain. Here we introduce a geochemical approach for estimating the flux from a vent field based on radon (²²²Rn) measurements in the overlying effluent plume. This method was applied successfully in September 1986 during a 23-day expedition to an active vent field on the 170-km Endeavour segment of the Juan de Fuca Ridge (Fig. 1a). We estimate the heat flux from this site to be $1-5 \times 10^9$ W.

The Endeavour segment of the Juan de Fuca Ridge has been studied intensively in recent years and is the focus of much ongoing and proposed research activity. In 1984, an active vent field at ~2,200 m depth within the axial valley (47°57' N, 129°06' W) was discovered and mapped with the *Alvin* submersible^{10,11}. The field covers 150 m × 250 m and contains many

'black smokers' venting fluids at temperatures up to 400 °C. A well-defined effluent plume exists at about 200 m above the depth of the vent field^{12,13}. Distinct excess radon levels have been observed at Endeavour both in vent water and in the effluent plume up to 17 km from the ridge axis¹⁴.

Radon is a chemically inert gas with a radioactive half life of 3.85 days. To estimate the hydrothermal heat flux from a vent field, we assume that the 'standing crop' of radon in the effluent (nonbuoyant) plume is at steady state, that is, we assume that the addition of radon to the plume from the vent field exactly equals the loss of radon from the plume due to radioactive decay. We assume that the flux of radon into the plume is constant over the duration of the study. The heat flux into the plume can be estimated by further assuming that the ratio of hydrothermal radon to heat in the plume directly above the vent field (the near-field) is constant over this same period. The heat content of the plume is calculated from the observed temperature anomaly (ΔT), which is defined as temperature in excess of that predicted from the background potential temperature-potential density relationship¹³. The effluent plume at Endeavour contains entrained colder water from deeper in the water column¹³, thus complicating the relationship between ΔT and hydrothermal heat¹⁵. However, salinity and temperature data can be used to quantify the effect of the entrained water on ΔT .

Our approach to estimating the heat flux from a vent field requires that we measure two parameters: the radon/ ΔT ratio in the near-field plume (before the plume has begun to disperse and/or age) and the standing crop of hydrothermally injected radon in the plume. The flux of chemical species enriched in hydrothermal fluids can be estimated in a similar manner by measuring the radon/species ratio in the near-field plume.

Water samples were collected at hydrocast sampling stations using a 12-bottle rosette and a conductivity-temperature-depth (CTD)/transmissometer probe. The samples were analysed for radon and other plume tracers, including ³He, manganese and methane. Temperature and light attenuation anomalies measured by towing the CTD package in a saw-tooth pattern through the deep water column allowed us to map the plume in real time. (Light attenuation indicates particle concentration). While the hydrocasts provided point samples for radon and other tracers, the 'tow-yos' allowed us to obtain an overall picture

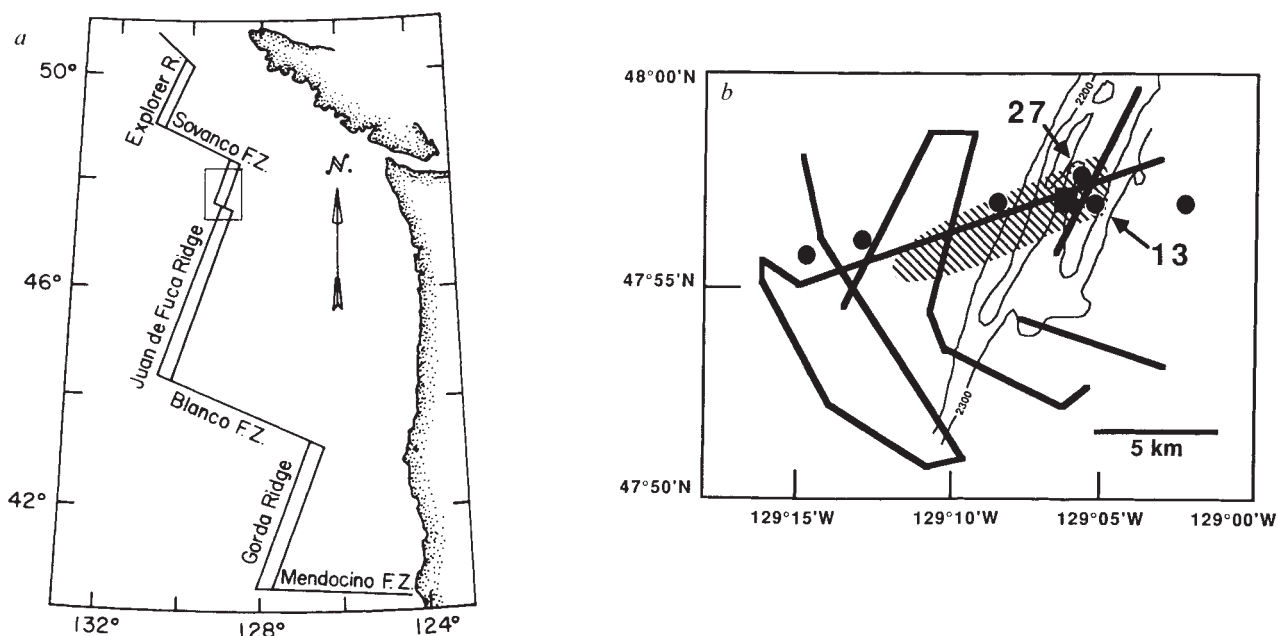


Fig. 1 *a*, Location of Endeavour segment of Juan de Fuca Ridge. *b*, Detailed map of flux study, 7–12 September 1986. The square indicates main active vent area (as confirmed by studies with the submersible *Alvin*^{10,11}). Contour lines (in metres) show ridge axis bathymetry. Heavy lines represent tow-yo tracks; circles represent hydrocast locations. The location of the near-field stations (13 and 27) used in our flux calculations are shown. Plume contours are based on temperature data from the tow-yos; the areas denoted with a dot pattern and a line pattern correspond to maximum temperature anomalies in the deep water column of 0 to ~0.04 °C and ~0.04 to ~0.05 °C, respectively.

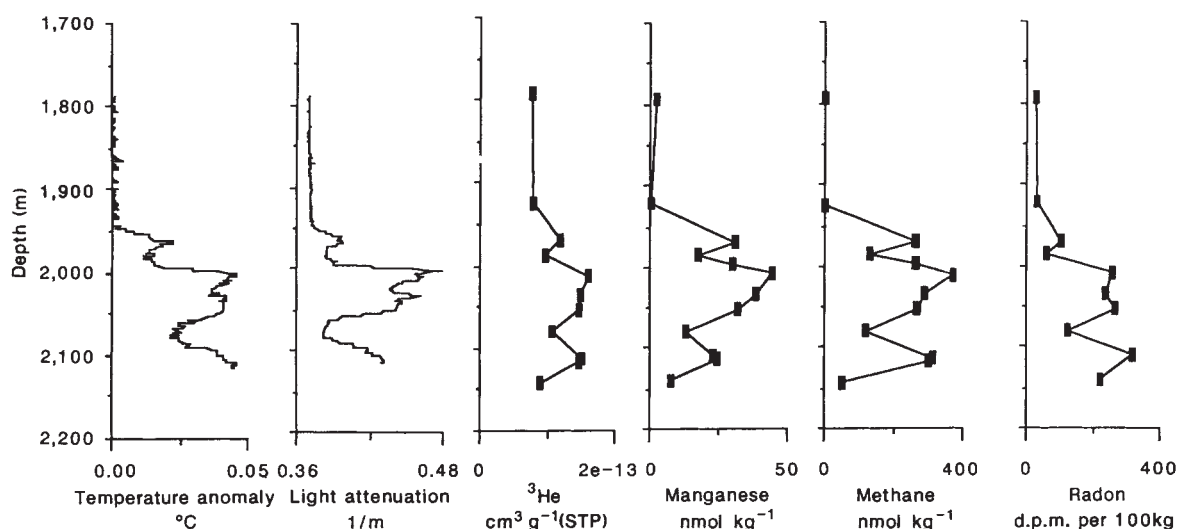


Fig. 2 Hydrocast profiles (temperature anomaly, light attenuation, ^3He , manganese, methane and radon) from Station 13 located in the near-field part of the plume (see Fig. 1). Temperature and light attenuation data below 2,115 m were not available. Background levels for ^3He and radon were $\sim 8 \times 10^{-14} \text{ cm}^3 \text{ g}^{-1} \text{ (STP)}$ and 30 d.p.m. per 100 kg respectively. The background radon levels are supported by decay of ^{226}Ra present in ambient ocean water. The correlation coefficients for the near-field (Stations 13 and 27) chemical data are Mn- CH_4 (0.92), Mn- ^3He (0.85), ^3He - CH_4 (0.96), Mn-Rn (0.28), CH_4 -Rn (0.43) and ^3He -Rn (0.52).

of plume characteristics.

The position of the plume varied during the 23-day study period. Our flux calculations were based on data from one 5-day study period which we believe to represent a 'snapshot' of the plume. The location of hydrocast stations and tow-yo tracks for this study are shown in Fig. 1*b*. The two near-field stations, 13 and 27, were occupied on the first and fifth day of our 5-day study. The fact that the radon levels at these two stations is nearly the same, and the radon/ ΔT and radon/species ratios agree within a factor of 2, gives us confidence in our assumption of a constant plume flux over the period in which our measurements were made.

A high degree of covariance was observed in the near-field for any two plume tracers, with the exception of radon (see Fig. 2). Below ~2,100 m the concentration of radon relative to the other tracers increases dramatically. This increase is seen clearly in Fig. 3, in which radon/ ^3He ratios for Stations 13 and 27 are plotted against depth.

The observed depth-dependence of the radon/ ^3He ratio can be explained by entrainment of radon-enriched bottom water by the rising (buoyant) plume. ^3He , a chemically inert non-radioactive gas, is a conservative tracer of hydrothermal activity and, unlike the measured ΔT signal, the concentration of ^3He in the effluent plume is a straightforward function of dilution

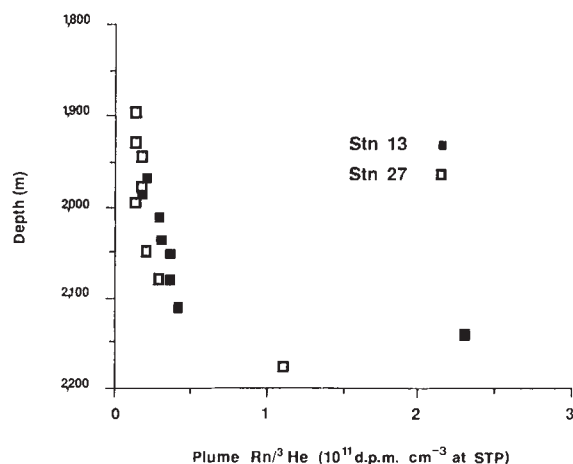


Fig. 3 Radon/ ^3He ratio versus depth for near-field stations, 13 and 27. Both radon and ^3He concentrations have been corrected for background levels.

of vent fluids. ^3He is removed from ocean crust by heated fluids. In contrast, radon can be extracted from ocean crust by circulating water of any temperature through alpha-recoil from the radioactive decay of its parent radium (^{226}Ra)^{16,17}. Low-temperature venting, which is ubiquitous in the area immediately surrounding the high-temperature vents, is the likely source of highly radon-enriched bottom water. In fact, low-temperature vents sampled at Endeavour have a much higher radon/temperature ratio than the high temperature vents¹⁷. A large percentage of the Endeavour effluent plume is entrained seawater from deeper in the water column¹³. The rising plume entrains surrounding seawater continuously as it ascends. Because the radon-enriched layer is near the bottom, we would expect, as observed, to find low radon/ ^3He ratios in the upper parts of the effluent layer and high radon/ ^3He ratios in the lower parts of the effluent layer.

For the purpose of estimating the hydrothermal heat flux, we chose to consider only the radon measurements from the water column shallower than $\sim 2,100$ m, at which depth the radon concentration sharply increases with respect to all other plume tracers. The average radon/ ΔT ratio at the centre of the plume ($\sim 2,000$ m) for the two near-field stations is ~ 0.04 d.p.m. $g^{-1} ^\circ\text{C}^{-1}$. (Radon is measured in d.p.m., disintegrations per minute). Assuming that the entrainment of deeper colder water contributes a negative temperature anomaly of $\sim 0.02 ^\circ\text{C}$ to the effluent plume, we chose ~ 0.03 d.p.m. $g^{-1} ^\circ\text{C}^{-1}$ as our initial radon/ ΔT value. We believe that this Rn/ ΔT ratio reflects more accurately the contribution of radon and heat from the actual vent field. This translates to an initial radon/heat ratio of ~ 55 atoms per joule.

Extrapolating the radon content of the entire plume from radon measurements at discrete hydrocast stations is complicated by the fact that the radon concentration in the water column is a function of both radioactive decay and dilution of the plume with ambient seawater. At each hydrocast location, the total plume radon content at depths of less than $2,100$ m was calculated by linear interpolation between data points. We used these radon values, combined with the tow-yo data, to estimate the standing crop of radon above $\sim 2,100$ m to be 8×10^{12} d.p.m. (see Fig. 4). To maintain the system (above $\sim 2,100$ m) in steady state, the vent field must be adding 8×10^{12} atoms of radon each minute.

Dividing the flux of radon into the plume (8×10^{12} atoms per minute) by the radon/heat ratio in the near-field plume (55 atoms per joule) gives a heat flux of $\sim 2.4 \times 10^9$ W into the plume at depths shallower than $2,100$ m. Based on the ^3He data, we

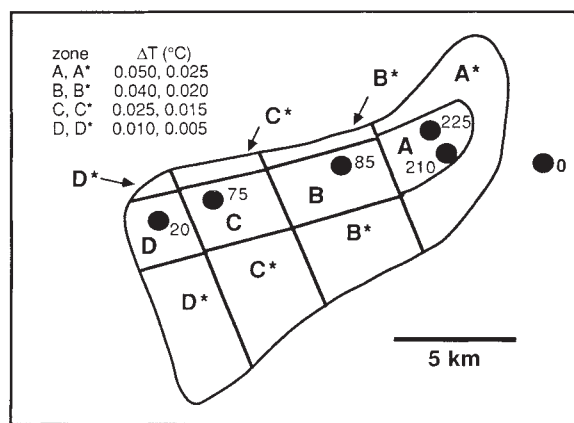


Fig. 4 Illustration of method used to estimate standing crop of radon in the plume. This figure shows the boundary of the plume and division of the plume into central zones (A, B, C, D) and outlying zones (A*, B*, C*, D*) as determined from temperature data from the tow-yos. Approximate ΔT values for each zone are given in the upper left corner of the figure. These temperature anomalies are maximum values from vertical profiles through the plume. The plume radon concentrations are only known at hydrocast locations, indicated by circles in the figure. The vertically integrated radon concentration above $2,100$ m (in units of 10^3 d.p.m. m^{-2}) is given for each station. We assumed that the radon content at each station was representative of the radon content of the section in which the station is located. We then used the temperature data to extrapolate to the radon content of the entire plume. The temperature anomalies in the outlying zones were observed to be a factor of 2 smaller than those in the corresponding central zones. Because lower temperature anomalies indicate dilution of the plume with ambient seawater, the radon levels in the outlying zones were assumed to be half those in the corresponding central zone. We estimated the total radon in each section by multiplying the radon level assumed for that section (d.p.m. m^{-2}) by the section area (m^2). Summing the contributions from all sections gives an estimate for the total standing crop of radon in the plume above $\sim 2,100$ m of 8×10^{12} d.p.m.

estimate that the plume above $\sim 2,100$ m contains $\sim 80\%$ of the total plume. Assuming that the percentage of the total heat flux represented by the flux into the plume above $2,100$ m is the same as the percentage of the total ^3He contained in the plume above this depth, we estimate the total heat flux from the Endeavour vent field to be $\sim 3.0 \times 10^9$ W. This heat flux is equivalent to the flux from 50 black smoker vents, or to the total hydrothermal heat flux from a 60 km medium-spreading ridge out to a width of one million years^{4,5}. It is difficult to assign error bars to our flux estimate; however, substitution of reasonable minimum and maximum values for the parameters used in this calculation suggests that the heat flux is $1-5 \times 10^9$ W. Measurements of ^3He , ^4He , manganese and methane in the near-field plume allowed us to estimate the vent flux for these species as $\sim 10^{-3}$ $\text{cm}^3 \text{ s}^{-1}$ (at standard temperature and pressure, STP), $\sim 10^2$ $\text{cm}^3 \text{ s}^{-1}$ (STP), $15-40$ g s^{-1} and $\sim 10^5$ $\text{cm}^3 \text{ s}^{-1}$ (STP) respectively. Using the approach of Lupton *et al.*¹³, we estimate that the volume flow rate of entrained water into the effluent plume is $1-3 \times 10^4$ $\text{m}^3 \text{ s}^{-1}$.

Our heat flux estimate agrees well with the calculation of $1-2 \times 10^9$ W by Baker and Massoth⁷ and is within the range of the more general estimates of Crane *et al.*⁸ for the Endeavour vent field. These two previous efforts relied heavily on assumptions about local oceanographic conditions (current strength and direction) which were difficult to constrain. Our estimate is free of these assumptions and therefore represents a second, completely independent, method of estimating the flux from a seafloor hydrothermal vent field.

We thank Captain W. Clappett and the crew of the R/V T. G. Thompson for assistance during expedition TT-204. We also

thank J. Fleishbein, D. Kovar, K. Thornberry, E. Olson and M. DeAngelis for sample analysis and data processing and D. Menzies and W. Rugh for valuable technical assistance at sea. This work was supported by the Division of Ocean Sciences of the NSF.

Received 15th April; accepted 14th July 1988.

1. Sclater, J. G., Jaupart, C. & Galson, D. *J. geophys. Res.* **86**, 11535–11552 (1981).
2. Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46**, 1–18 (1979).
3. Hekinian, R. *et al. Science* **207**, 1433–1444 (1980).
4. MacDonald, K. C., Becker, K. Spiess, F. N. & Ballard, R. D. *Earth planet. Sci. Lett.* **48**, 1–7 (1980).
5. MacDonald, K. C. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. *et al.*) 27–51 (Plenum, New York, 1983).
6. Converse, D. R., Holland, H. D. & Edmond, J. M. *Earth planet. Sci. Lett.* **69**, 159–175 (1984).
7. Baker, E. T. & Massoth, G. J. *Earth planet. Sci. Lett.* **85**, 59–73 (1987).
8. Crane, K. C. *et al. J. geophys. Res.* **90**, 727–744 (1985).
9. Little, S. A., Stolzenbach, K. D. & von Herzen, R. P. *J. geophys. Res.* **92**, 2587–2596 (1987).
10. MERGE Group. *EOS* **65**, 1111 (1984).
11. Hammond, S. R. *et al. EOS* **65**, 1111 (1984).
12. Lupton, J., Delaney, J., Johnson, H. P. & Tivey, M. *EOS* **65**, 975 (1984); **66**, 929 (1985).
13. Lupton, J., Delaney, J., Johnson, H. P. & Tivey, M. *Nature* **315**, 621–623 (1985).
14. Kadko, D. & Lupton, J. E. *EOS* **65**, 974 (1984).
15. Speer, K. G. thesis, WHOI-MIT (1988).
16. Dymond, J., Cobler, R., Gordon, L., Biscaye, P. & Mathieu, G. *Earth planet. Sci. Lett.* **694**, 417–429 (1983).
17. Kadko, D. & Moore, W. *Geochim. cosmochim. Acta* **52**, 659–668 (1988).

Normal potassium, inherited argon in Zaire cubic diamonds

F. A. Podosek*, J. Pier*, O. Nitoh†, S. Zashu‡ & M. Ozima‡

* Department of Earth and Planetary Sciences and McDonnell Center for the Space Sciences, Washington University, St Louis, Missouri 63130, USA

† Faculty of Engineering, Tokyo University of Agriculture and Technology, Tokyo 184, Japan

‡ Geophysical Institute, University of Tokyo, Tokyo 113, Japan

A group of 10 cubic diamonds from Zaire has been found¹ to contain correlated concentrations of ⁴⁰Ar and K which, interpreted as a whole-rock K–Ar isochron with the usual assumptions, yield the unreasonable age of 6.0 Gyr. The same age has also been determined² by ⁴⁰Ar–³⁹Ar analysis of four additional diamonds from the same group. One explanation for these data is that the potassium in these diamonds is isotopically anomalous, specifically, enriched at least twofold in ⁴⁰K relative to normal K, as might have resulted from preservation of cosmic-ray-induced spallation K produced in iron-rich planetesimals before accretion to the Earth. We have examined the isotopic composition of K in three diamonds from the same group, and report here that, within analytical uncertainties (about 1%), ⁴⁰K is present in normal abundance, and the hypothesis of isotopically anomalous K is not supported. An alternative explanation is that the ⁴⁰Ar in these diamonds is a trapped or 'excess' component, but this would require the unusual circumstance of a good correlation between ⁴⁰Ar and K (for example, by incorporation of a fluid phase), but a lack of correlation between ⁴⁰Ar and other noble gas species, including ³⁶Ar.

The diamonds in the group studied, commonly described as 'cuboid', are opaque white to yellowish microcrystalline aggregates. This group of samples, generally ~3 mm on edge and weighing ~0.2–0.3 g, was obtained commercially, and it is not possible to specify their provenance more precisely than being in Zaire.

Zashu *et al.*¹ studied 10 diamonds from this group, measuring K by instrumental neutron activation analysis (INAA) and Ar by noble gas mass spectrometry. They found K concentrations ranging from 10 to 47 p.p.m. and ⁴⁰Ar concentrations proportional to K concentrations in the ratio 0.177 cm³ STP g⁻¹; by the

usual assumptions involved in the K–Ar method, this ratio corresponds to an age of 6.0±0.3 Gyr. Takigami *et al.*² found apparent ages in the range 5.3–5.8 Gyr in applying the ⁴⁰Ar–³⁹Ar technique to four additional diamonds from the same batch.

These unreasonable ages led these authors^{1,2} to question the assumptions involved in K–Ar dating. Specifically, the isotopes used as tracers for K are ⁴¹K (for INAA) or ³⁹K (for the ⁴⁰Ar–³⁹Ar method), whereas it is the radioactive isotope ⁴⁰K that is the parent of ⁴⁰Ar, so that in both cases it must be assumed that ⁴⁰K is present in normal proportion to the major isotopes of K. The anomalously high ages might thus be attributable to an anomalously high abundance of the scarce isotope ⁴⁰K, relative to the major isotopes. The effect would have to be quite large: more than a factor of two enrichment of ⁴⁰K even if the actual age of the diamonds were 4.6 Gyr, and an additional factor of two for every ⁴⁰K-half-life (1.25 Gyr) by which the actual age of the diamonds is less than 4.6 Gyr. Zashu *et al.*¹ speculated that ⁴⁰K enrichment might have been generated by cosmic-ray-induced spallation reactions on Fe+Ni in K-poor metal planetesimals before their accretion to the Earth, an effect that is readily observed in iron meteorites³. The diamonds would have to have formed early in Earth history in a region where such isotopically anomalous K was being exsolved from the metal, before this K could mix with normal terrestrial K. This particular cosmochemical scenario or, more generally, isotopically anomalous K however generated, seems a rather remote possibility, but one which would have profound implications if true, and the present study of direct measurement of ⁴⁰K abundance was undertaken to examine this possibility.

We have examined K in three additional Zaire diamonds from the same group as in the prior studies^{1,2}, and, for comparison, from one South African diamond. Diamond samples were combusted in an externally heated quartz tube with a flow of O₂ gas over the sample. After the combustion, the walls and bottom of the quartz tube were washed in HCl, and the resultant solution designated the 'S' sample. For two of the Zaire diamonds, gas exhausted during the combustion was bubbled through water in another vessel, and the resultant solution designated the 'T' sample. Small aliquots of the samples were used for mass spectrometric analysis of K (the rest being reserved for other studies) by the 'direct loading' approach, that is, with no further chemical processing such as ion-exchange purification. For the two Zaire diamonds for which both S and T samples were taken, separate small aliquots of sample solution were spiked with a ⁴¹K tracer solution in order to determine K abundances by isotopic dilution.

Isotopic analysis was performed on a VG-354 thermal emission mass spectrometer operated in single-collector Faraday cup mode. In unspiked runs, beam intensity was 4×10⁻¹¹ A of ³⁹K and measured ⁴⁰K/³⁹K ratios were corrected for instrumental discrimination by assuming normal composition (Table 1) for ⁴¹K/³⁹K. The precision-limiting factor for ⁴⁰K/³⁹K in these analyses is electronic noise in the detection system. In

Table 1 Results for potassium in diamonds

Sample		Wt (mg)	[K] (p.p.m.)	Δ ₄₀ * (%)
Zaire 8601	S	120		-0.1±1.8
Zaire 8801	S	313	6.9	0.9±0.8
	T		2.9	0.0±1.1
Zaire 8802	S	305	18.9	0.0±1.1
	T		16.5	0.2±0.8
S. Africa 8602	S	116		-0.9±1.2

* Measured in unspiked analyses, and stated as per cent deviation of sample ⁴⁰K/³⁹K from mean value in calibrations (which is 0.6% higher than the generally accepted value ⁴⁰K/³⁹K = 0.00012516). Corrections for instrumental discrimination based on ⁴¹K/³⁹K = 0.072168. Standard values for K composition from refs 4 and 5.

spiked runs the beam intensity was 1×10^{-11} A of ^{39}K and only $^{41}\text{K}/^{39}\text{K}$ was measured; this ratio was corrected for instrumental discrimination by assuming a discrimination of 11 per mil per AMU, based on unspiked K calibrations. The precision-limiting factor in the spike analyses is reproducibility of discrimination, and we estimate that the uncertainty of the isotope dilution data (Table 1) is $\leq 2\%$.

We monitored masses 38–44 AMU during these analyses, but observed no peaks other than those of K. Although Ca, among other elements, was loaded on the filament along with K, it did not generate a detectable ion beam at the low filament temperatures used for K analysis, and no interference corrections were applied.

Mass spectrometry was calibrated by analysis of reagent K and K from a natural sylvite by the same procedures. $^{40}\text{K}/^{39}\text{K}$ calibration measurements were reproducible at the level indicated by in-run statistics and averaged 0.6% higher than the generally accepted^{4,5} ratio ($^{40}\text{K}/^{39}\text{K} = 0.00012516$). We do not consider this as significantly different from the accepted normal composition but rather as a laboratory bias effect, which could be accounted for in terms of ion scattering into the baseline or imperfect correction for finite amplifier decay time, etc. Sample $^{40}\text{K}/^{39}\text{K}$ ratios in Table 1 are presented as percentage deviations from the mean $^{40}\text{K}/^{39}\text{K}$ ratio in our calibration analyses.

Reagent and procedural blank analyses were also performed. The largest blank observed was ~ 10 ng of K. The smallest K sample (for which we have isotope dilution results) was ~ 450 ng, so the blanks are small compared to the samples. It is not possible to determine whether K was extracted and collected with complete efficiency, but total observed concentrations for 8801 and 8802 (Table 1) are within the range previously observed by INAA¹. We are thus confident that the isotopic data in Table 1 reflect the K in the diamonds rather than an experimental artifact.

As seen in Table 1, all the $^{40}\text{K}/^{39}\text{K}$ ratios observed are normal within errors. It should also be noted that within comparable limits the $^{41}\text{K}/^{39}\text{K}$ ratios are also normal, that is, that the inferred discriminations during sample analyses were comparable to inferred discriminations during calibration analyses. K in these diamonds is thus not fractionated (within a few per mil per AMU) relative to normal K.

The Zaire diamonds in the present study, of course, are not the same physical specimens as those which have yielded 6 Gyr apparent ages. It is not plausible, however, that 14 diamonds selected for Ar analysis would all have a several-fold enrichment of ^{40}K while three others, from the same group, selected for ^{40}K analysis would show no excess at the level of 1%. We conclude that the high apparent ages^{1,2} in Zaire diamonds cannot be attributed to anomalous enrichment of ^{40}K .

The present data also rule out a variant hypothesis^{1,2} that excessive K–Ar ages might be attributable to a pressure-induced acceleration of the electron-capture decay branch of ^{40}K . This would result in a deficiency of present ^{40}K abundance, for example, by 35% in order to cause 4.6 Gyr samples appear to be 6.0 Gyr old, and by a larger fraction if the true age is less than 4.6 Gyr. This is not consistent with the observed data. We dismiss the possibility of cosmogenic enhancement of ^{40}K coupled with accelerated electron-capture decay in just such proportions to coincidentally result in normally abundant ^{40}K .

The present results do not modify or extend the previous observations^{1,2}, of course, but only eliminate interpretations which involve isotopically abnormal K. Viable interpretations can be divided into two groups: either the observed ^{40}Ar is the *in situ* decay product of ^{40}K , or it is not.

If the ^{40}Ar was produced *in situ*, that is, the diamonds really are 6.0 Gyr old, it is difficult to advance an unspectacular hypothesis. Discounting overturn of the overwhelming evidence for a 4.6 Gyr age for the Solar System, these diamonds, or at least parts of them, would have to be presolar grains that were never mixed with the bulk of Earth materials. Presolar diamonds

have indeed been observed in primitive meteorites⁶, and it could be speculated that presolar diamonds constitute or served as nucleation sites for the Zaire diamonds, never mixing with other Earth materials. In such case, however, it would be expected that other elements carried in these diamonds would display nearly endemic isotopic anomalies. By hypothesis, these 'other elements' would necessarily include most of the K. As K is not anomalous, it is difficult to defend the plausibility of an *in situ* interpretation. It should be stressed, however, that by normal criteria—correlation of ^{40}Ar with K, from one sample to another to form a whole-rock isochron, and in stepwise heating in ^{40}Ar – ^{39}Ar analysis—the observed ^{40}Ar would generally be interpreted as *in situ* if the resultant age were reasonable; it is the unreasonable age that makes the *in situ* interpretation implausible.

^{40}Ar that is not generated *in situ* is an inherited, or 'excess', or 'trapped' component. This simply means that it was present at the time of isotopic closure of the sample in question, and trapped noble gases, especially excess ^{40}Ar , are well known to occur in a variety of volcanic rocks^{7–10}, notably quenched MORBs (mid-ocean ridge basalts). Although it is not difficult, *a priori*, to accept the possibility of excess ^{40}Ar in the Zaire or other diamonds, if the ^{40}Ar in the Zaire diamonds is indeed a trapped component it has some noteworthy features which provide significant constraints on its and the diamonds' origin.

Foremost of these is the close correlation with K, a feature that has not previously been noted for excess ^{40}Ar . It is difficult to stipulate how this correlation could have arisen and been maintained unless both the ^{40}Ar and the K were transported and entered the diamonds in a fluid. Secondly, although ^{40}Ar in these diamonds correlates with K it does not correlate with other noble gases^{1,2}, specifically ^{36}Ar . This requires that the other gases were introduced into the diamonds by a different mechanism than the $^{40}\text{Ar} + \text{K}$. It might be supposed that this 'different mechanism' is only contamination with atmospheric gases, but the thermal release patterns and isotopic structure of Ne^{11} argue against such a simple interpretation. Finally, the observed $^{40}\text{Ar}/^{36}\text{Ar}$ ratios are very high, in some cases in the range 20,000–30,000. Most of the ^{40}Ar —nearly all of it if the true age of the diamonds is significantly less than 4.6 Gyr—would be trapped rather than *in situ*. The trapped Ar in these Zaire diamonds would thus have higher $^{40}\text{Ar}/^{36}\text{Ar}$ than is generally found in MORBs (with rare exceptions⁸ MORB observations show $^{40}\text{Ar}/^{36}\text{Ar} < 15,000$, the value which Fisher¹⁰ proposes as characteristic of the MORB source). In most models of Ar evolution⁹ this would correspond to more rapid or more extensive 'depletion' of the source region of these diamonds, in comparison with the source of MORBs.

Although there are a number of noble gas observations in other diamonds¹², it is difficult to determine the existence or character of any excess ^{40}Ar , as this requires both K concentrations and independent chronological constraints. There is at least one set of data¹², however, that provides an interesting contrast to the Zaire diamonds, in that the close correlation of ^{40}Ar with ^{36}Ar in stepwise release from an Australian diamond is strongly suggestive of a trapped Ar component, one with a very low $^{40}\text{Ar}/^{36}\text{Ar}$ ratio, ~ 320 . Ar evolution models⁹ indicate that such a low ratio corresponds to a relatively 'undepleted' mantle reservoir. It may be that diamonds exhibit noble gas isotopic structures, particularly for Ar, which parallel those of more common igneous rocks, with some, for example, the Australian diamond¹², originating in relatively undepleted or less-depleted reservoirs analogous to the sources of some ocean-island volcanics, and others, the Zaire diamonds for example, originating in more depleted reservoirs analogous to the MORB source.

Received 27 May; accepted 1 July 1988.

1. Zashu, S., Ozima, M. & Nitoh, O. *Nature* **323**, 710–712 (1986).
2. Takigami, Y., Ozima, M. & Zashu, S. *Eos, Wash.* **69**, 503 (1988).

3. Imamura, M., Shima, M. & Honda, M. *Z. Naturf.* **35A**, 267-279 (1980).
4. Steiger, R. H. & Jager, E. *Earth planet. Sci. Lett.* **36**, 359-362 (1977).
5. Garner, E. L., Murphy, T. J., Gramlich, J. W., Paulsen, P. J. & Barnes, I. L. *J. Res. natn. Bur. Stand.* **79A**, 713-725 (1975).
6. Lewis, R. S., Tang, M., Wacker, J. F., Anders, E. & Steel, E., *Nature* **326**, 160-162 (1987).
7. Funkhouser, J. G., Fisher, D. E. & Bonatti, E. *Earth planet. Sci. Lett.* **5**, 95-100 (1968).
8. Allègre, C. J., Staudacher, T., Sarda, P. & Kurz, M. *Nature* **303**, 762-766 (1983).
9. Ozima, M. & Podosek, F. *Noble Gas Geochemistry* (Cambridge University Press, 1983).
10. Fisher, D. E. *Geochim. cosmochim. Acta* **51**, 2531-2541 (1987).
11. Ozima, M. & Zashu, S. *Geochim. cosmochim. Acta* **52**, 19-25 (1988).
12. Ozima, M. A. *rev. Earth planet. Sci.* (in the press).

Submarine hot springs and the origin of life

Stanley L. Miller* & Jeffrey L. Bada†

* Department of Chemistry (B-017), University of California at San Diego, La Jolla, California 92093, USA

† Amino Acid Dating Laboratory (A-012B), Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, USA

The discovery of hydrothermal vents at oceanic ridge crests and the appreciation of their importance in the element balance of the oceans is one of the main recent advances in marine geochemistry¹. It is likely that vents were present in the oceans of the primitive Earth because the process of hydrothermal circulation probably began early in the Earth's history². Here we examine the popular hypothesis³⁻⁷ that life arose in these vents. This proposal, however, is based on a number of misunderstandings concerning the organic chemistry involved. An example is the suggestion that organic compounds were destroyed on the surface of the early Earth by the impact of asteroids and comets, but at the same time assuming that organic syntheses can occur in hydrothermal vents⁷. The high temperatures in the vents would not allow synthesis of organic compounds, but would decompose them, unless the exposure time at vent temperatures was short⁸⁻¹¹. Even if the essential organic molecules were available in the hot hydrothermal waters, the subsequent steps of polymerization and the conversion of these polymers into the first organisms would not occur as the vent waters were quenched to the colder temperatures of the primitive oceans.

The hydrothermal-vent origin-of-life hypothesis can be divided into three steps: (1) synthesis of amino acids and other essential organic compounds from the dissolved methane and other gases, initially at high temperatures, as they pass through a temperature gradient (>350° to ~2°C); (2) synthesis of peptides and nucleotides by thermal dehydration and other high-temperature reactions; and (3) synthesis of protocells or RNA-like molecules at some stage in the thermal gradient, and the conversion of these protolife forms into living organisms as the hydrothermal waters emerge from the vents. We will evaluate the plausibility of each of these steps.

An important consideration is how long organic compounds in vent waters are exposed to >350°C and how rapidly these hot hydrothermal solutions are quenched to ambient seawater temperatures. At spreading ridges, seawater is heated to temperatures in excess of 350°C. The estimated residence time for seawater at 350°C is 22-45 yr¹². The transit time from this zone of hot water-basalt interactions to the vent seawater interface is <1 yr¹³, but there is no thermal gradient until the waters at ~350°C exit the vents and are rapidly quenched to deep ocean temperatures. Thus, the only thermal gradient in the system is in the area near the vent opening, which is where the proposed organic synthesis and further steps in the origin of life would have to take place.

Step 1. The pyrolytic synthesis of amino acids is not an efficient process. An example is the reported synthesis of most protein amino acids by the pyrolysis of CH₄, NH₃ and H₂O at 900-1,100°C for a few seconds in the presence of various catalysts¹⁴.

It was subsequently shown that the yields were extremely low (0.007%) and that the products were mostly β-amino acids, not the α-amino acids found in proteins¹⁵.

Another type of pyrolytic synthesis is the Fischer-Tropsch reaction¹⁶. Gaseous mixtures of CO and H₂, which are at equilibrium at elevated temperatures (>800°C), become unstable at lower temperatures (200-400°C) and many organic compounds can be synthesized with appropriate catalysts¹⁷, especially in the presence of NH₃. This type of process may have been important in the Solar Nebula. The products are produced in fair yield and are quite diverse, and include many amino acids, purines and pyrimidines. The essential feature of this synthesis is the short contact time of the gases with the catalysts at 200°-400°C; long contact times decompose the products¹⁰. In order for this process to occur in the vents, a narrow temperature-time window with the waters in contact with the catalysts would be required, or otherwise decomposition would predominate over synthesis. Moreover, the Fischer-Tropsch synthesis has never been performed in an aqueous environment, and the H₂S present in vent waters¹ would poison the Fischer-Tropsch catalysts¹⁷.

High-temperature origin-of-life theories must deal with the general problem that at high temperatures organic compounds are highly unstable and decompose rapidly. The available stability data have been reviewed⁸⁻¹¹ and show that the essential molecules would have short lifetimes at temperatures above 200°C. The high pressures (200-400 atm) associated with deep ocean hydrothermal systems would not greatly retard the decomposition rates^{9,11}. Even assuming a large volume of activation of 50 ml mol⁻¹, a pressure of 400 atm would decrease the rate of decomposition by a factor of only 2.5.

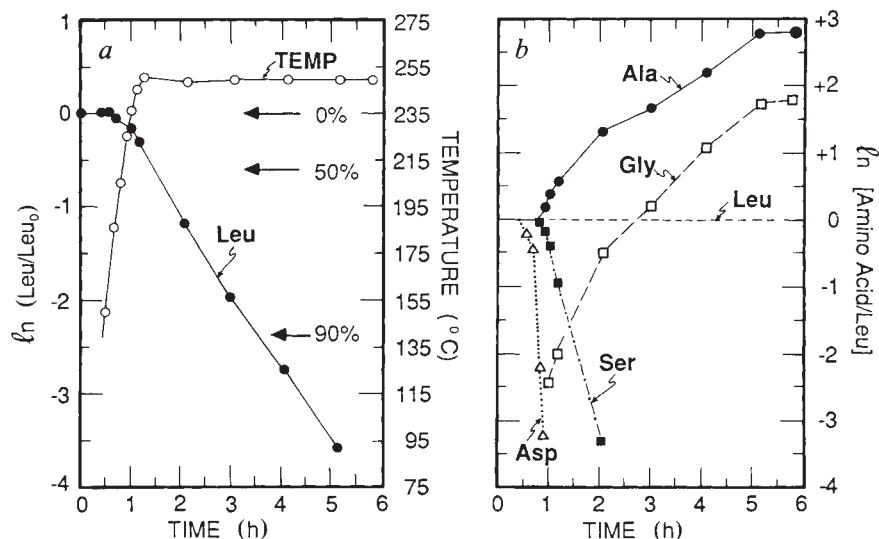
To demonstrate further that biomolecules are very unstable at high temperatures and that their stability is not enhanced by high pressures, we have carried out several experiments at 250°C and 265 atm at neutral pH. Figure 1a shows the rate of decomposition of leucine. Substantial decomposition had taken place even before 250°C was attained. The estimated half life for leucine decomposition at 250° and 265 atm is ~15-20 min. This half life at 250°C is less than that estimated from earlier measurements¹⁸ at the vapour pressure of water (40 atm) using unbuffered solutions. Figure 1b gives the amounts relative to leucine for an amino acid mixture that originally contained equal molar amounts of aspartic acid, serine, alanine and leucine; glycine was not initially present. Aspartic acid and serine were found to decompose much more rapidly than leucine; serine has a half life of only a few minutes or less, whereas the aspartic acid value is <1 min. These values are consistent with extrapolated estimates⁹. Alanine appears to be more stable than leucine, and glycine was produced during the course of the heating experiment. The glycine produced comes from the aldol cleavage of serine and the rise in the relative amount of alanine is due to its synthesis from serine by dehydration and reduction reactions¹⁹.

Sugars would survive at 250°C for seconds at most^{20,21}. The hydrolysis of HCN, an important component in many prebiotic syntheses, would also be very rapid. At 250°C, pH=8, the extrapolated half life for HCN hydrolysis to formamide is 3 s; at pH=4, the half life is 3 h²². The corresponding values at 350°C are 0.15 s and 16 min respectively.

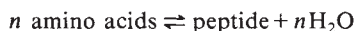
There is the additional problem of maintaining molecular chirality at high temperatures²³. Modern organisms require a high degree of chiral integrity in their biomolecules. Our studies with isoleucine at 250°C and 265 atm indicate that the epimerization rate of this amino acid to alloisoleucine exceeds the decomposition rate.

On this basis, we would predict that organic synthesis would not occur in hydrothermal vent waters, which should thus be devoid of organic molecules. This can be tested using highly sensitive analytical methods, such as laser-based fluorescence detection, to survey vent waters. It should be emphasized that

Fig. 1 Decomposition of amino acids at 250 °C and 265 atm. The experimental system, which has been described elsewhere³⁵, allowed periodic sampling without cooling or depressurizing. The amino acids were dissolved in 0.01 M HEPES (N-2-hydroxyethylpiperazine-N-2 ethane sulphonic acid) buffer, initial pH of 7.0. The pH measured after heating for 5 h was 7.9. The sample mixture contained 10^{-3} M aspartic acid, serine, alanine and leucine; the ionic strength was adjusted to 0.7 M with NaCl. The amino acid concentrations were determined on a Beckman Model 119 amino acid analyser interfaced with a DEC PDP-11 computer used for data acquisition and analysis. *a*, The decomposition of leucine plotted in the form of irreversible first-order kinetics, and the temperature measured during the course of the experiment. The arrows indicate various percentages of decomposition. *b*, The amounts relative to leucine of the indicated amino acids. Glycine was not initially present, but was produced from serine during the heating experiment (see text).



important organic geochemical reactions do occur as hot vent waters penetrate the nearby sediments. For example, petroleum-like material has been found at some vent localities²⁴. But this is not an example of abiotic synthesis in vent waters; rather the petroleum results from the decomposition, alteration and catagenesis of biologically derived sedimentary organic matter. **Step 2.** The polymerization of the amino acids into peptides will not work under vent conditions as the equilibrium



is unfavourable in the presence of liquid water at all temperatures²⁵. Polypeptides have been reported to be synthesized in experiments carried out between 180 °C and 200 °C under dry conditions²⁶, where the equilibrium is shifted to the right because of the low activity (relative humidity) of water. It would obviously be impossible to have dry conditions in a submarine vent. The use of compounds such as hydrogen cyanide and cyanamide, for example, which have high free energies of hydrolysis, makes the thermodynamics more favourable, but it remains to be shown whether such reactions will in fact take place under the vent conditions of high temperatures and dilute solutions.

Even if some efficient mechanism for peptide synthesis existed, these peptides would be rapidly hydrolysed back to free amino acids at the high temperatures of the vents. At 250 °C, various alanine peptides⁹ were hydrolysed in only a few minutes, and polyglycine was completely hydrolysed in 6 h¹¹. We attempted to measure the rate of hydrolysis of bovine serum albumin at 250 °C and 265 atm (see Fig. 1 for experimental conditions) but the rate is so fast that none of the original protein or any large peptides could be detected by gel electrophoresis after the temperature had risen to only 140 °C.

Efficient synthetic pathways yielding deoxy- and ribonucleotides would also be difficult to imagine in a high-temperature environment. The prebiotic synthesis of purine nucleosides by heating the bases with sugars is inefficient and was carried out under dry conditions²⁷. As noted previously, sugars are very unstable at elevated temperatures and high-energy phosphates are hydrolysed rapidly at 250 °C (for example, the half life for hydrolysis⁹ of ATP is <1 s).

Step 3. The proposed protocells of Corliss *et al.*³ are undefined. If they envision microspheres of Fox²⁶, then we can say that these are poor precursors to a living organism. No one has ever claimed that Fox's microspheres are living, and it is generally agreed that the microspheres were not on the evolutionary pathway to the first living organisms.

The hypothesis that the first primitive organisms were made up of RNA-like molecules has become popular since the dis-

covery of self-splicing in RNA²⁸. This concept is very attractive for a number of reasons^{29,30}, although it is likely that there were RNA-like precursors that did not contain ribose³¹. A hydrothermal vent environment, however, is incompatible with the RNA hypothesis because of the inherent instability of these molecules at high temperatures. For example, extrapolation of the rates of hydrolysis^{8,32} of RNA at pH 7 gives a half life of 2 min at 250 °C for the hydrolysis of each phosphodiester bond; at 350 °C the half life is 4 s. For DNA, the half lives for depurination³² of each nucleotide at pH 7 are nearly the same as the hydrolysis rates of RNA. At the pH 4 of vent waters, these half lives would be shortened by a factor of $\sim 10^3$. Claims that bacteria can be cultured from hydrothermal vent waters³³ has been cited as evidence that the fragile molecules associated with living organisms can survive at 250 °C. But what was reported to be bacterial growth at 250 °C was subsequently demonstrated to be an artefact that arose from heating the culture medium used in the experiment³⁴.

In summary, we have shown that the proposal for a hydrothermal-vent origin of life fails each of the three proposed steps involved in the origin of life. Any origin-of-life theory that proposes conditions of temperature and time inconsistent with the stability of the compounds involved can be dismissed solely on this basis, unless some protective mechanism exists, but no such mechanisms are known at present. This is not to say that the hydrothermal vents were a negligible factor on the primitive Earth. They would have had a function in the elemental balance of the oceans as they do today. The vents would have been important in the destruction rather than in the synthesis of organic compounds in the primitive oceans, because the entire ocean currently passes through the vents in 10^7 yr (ref. 1). In fact, this latter process was an important factor in fixing the upper limit for the organic compound concentration in the oceans of the primitive Earth²².

We thank A. A. Yayanos, R. Van Bortel, J. Trent and M. Boehm for assistance with the heating experiment. This research was supported by grants from NASA (S.L.M.) and the California Space Institute, University of California (J.L.B.).

Received 22 March; accepted 24 June 1988.

1. Edmond, J. M. *et al.* *Nature* **297**, 187-191 (1982).
2. Sleep, N. H. *Origins of Life* **16**, 179-180 (1986).
3. Corliss, J. B., Baross, J. A. & Hoffman, S. E. *Oceanologica Acta* **4** suppl. 59-69 (1981).
4. Baross, J. A. & Hoffman, S. E. *Origins of Life* **15**, 327-345 (1985).
5. Nisbet, E. G. *The Young Earth* (Allen & Unwin, Winchester, Massachusetts, 1987).
6. Knoll, A. H. *Science* **239**, 199-200 (1988).
7. Maher, K. A. & Stevenson, D. J. *Nature* **331**, 612-614 (1988).
8. Miller, S. L. & Orgel, L. E. *The Origins of Life on the Earth* 118-128 (Prentice-Hall, Englewood Cliffs, 1974).

9. White, R. H. *Nature* **310**, 430–432 (1984).
10. Miller, S. L., Urey, H. C. & Oro, J. *J. molec. Evol.* **9**, 59–72 (1976).
11. Bernhardt, G., Lüdermann, H.-D., Jaenicke, R., König, H. & Stetter, K. O. *Naturwissenschaften* **71**, 583–586 (1984).
12. Turekian, K. K. & Cochran, J. K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6241–6244 (1986).
13. Macdonald, K. C., Becker, K., Spiess, F. N. & Ballard, R. D. *Earth planet. Sci. Lett.* **48**, 1–7 (1980).
14. Harada, K. & Fox, S. W. *Nature* **201**, 335–336 (1964).
15. Lawless, J. G. & Boynton, C. G. *Nature* **243**, 405–407 (1973).
16. Storch, H. H., Golumbic, N. & Anderson, R. B. *The Fischer-Tropsch and Related Syntheses* (Wiley, New York, 1951).
17. Hayatsu, R. & Anders, E. *Top. curr. Chem.* **99**, 1–37 (1981).
18. Vallentyne, J. R. *Geochim. cosmochim. Acta* **32**, 1853–1856 (1968).
19. Bada, J. L., Shou, M.-Y., Man, E. H. & Schroeder, R. A. *Earth planet. Sci. Lett.* **41**, 67–76 (1978).
20. Feather, M. S. & Harris, J. F. *Adv. Carbohydr. Chem. Biochem.* **28**, 161–224 (1973).
21. Mopper, K., Dawson, R., Leibzeit, G. & Ittekkot, V. *Mar. Chem.* **10**, 55–66 (1980).
22. Stribling, R. & Miller, S. L. *Origins of Life* **17**, 261–273 (1988).
23. Bada, J. L. & Miller, S. L. *BioSystems* **20**, 21–26 (1987).
24. Simoneit, B. R. T. *Can. J. Earth. Sci.* **22**, 1919–1929 (1985).
25. Meggy, A. B. *J. chem. Soc.* **1956**, 1444–1454 (1956).
26. Fox, S. & Dose, K. *Molecular Evolution and the Origin of Life* (Freeman, San Francisco, 1972).
27. Fuller, W. D., Sanchez, R. A. & Orgel, L. E. *J. molec. Evol.* **1**, 249–257 (1972).
28. Cech, T. R. & Bass, B. L. *A. Rev. Biochem.* **55**, 599–629 (1986).
29. Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4360–4363 (1986).
30. Westheimer, F. H. *Nature* **319**, 534–536 (1986).
31. Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4398–4402 (1987).
32. Ginoza, W., Hoelle, C. J., Vessey, K. B. & Carmack, C. *Nature* **203**, 606–609 (1964).
33. Baross, J. A. & Deming, J. W. *Nature* **303**, 423–426 (1983).
34. Trent, J. D., Chastain, R. A. & Yayanos, A. A. *Nature* **307**, 737–740 (1984).
35. Yayanos, A. A., Van Bostel, R. & Dietz, A. *Appl. envir. Microbiol.* **46**, 1357–1363 (1983).

A novel level of interactions in plant–insect systems

David Ng

Department of Zoology, University of Texas, Austin, Texas 78712, USA

The importance of variation in the ability of individuals to survive or perform on different types of resource (implicit in evolutionary studies) has been discussed in both host–parasite¹ and plant–insect systems^{2–4}. But variation in host-choice behaviour and its interaction with offspring performance on variable resources⁵ has been studied only in insect populations that each use several plant species^{6,7}. Here I show that in a natural population of the butterfly *Euphydryas editha* (Nymphalidae), some females are specialized in both oviposition behaviour and offspring performance. These specialists prefer to lay eggs on certain individuals of the host plant *Pedicularis semibarbata* (Scrophulariaceae) that also improve the survival of their offspring. Other *E. editha* females in the same population are not so specialized. They show no oviposition preference among individual *P. semibarbata*, and survival of their offspring is not correlated with the plant suitability categories defined by the specialists. Thus information vital to understanding the true nature of the interactions in a plant–insect system is lost when the process is described using either the average ability of an insect population to exploit a plant population or the average ability of a plant to resist attack by an insect species.

The study site is located at Generals Highway, Sequoia National Forest, Tulare County, California^{8,9}, where adult *E. editha* fly during a four-week period about four weeks after snowmelt. Female *E. editha* were captured randomly and their oviposition behaviour determined by using an oviposition preference test^{10,11} in which sequences of phenologically and morphologically similar, naturally growing *P. semibarbata* were used. When a female responded to the plant on which it was placed by crawling to the base and curling its abdomen with ovipositor extruded in an attempt to lay eggs, it was recorded as having ‘accepted’ that plant. Failure to accept a plant after 5 min of exposure was recorded as a rejection. The acceptance

or rejection of a plant is repeatable and not affected by the order in which individual plants are offered¹⁰.

Twelve groups of 5–7 females were each tested with a different sequence of 4–6 plants. Two behavioural categories were found: (1) non-discriminating (accepting all *P. semibarbata* individuals offered); (2) discriminating (repeatedly accepting certain *P. semibarbata* individuals but consistently rejecting others). The females were classified as non-discriminators or discriminators accordingly. The *P. semibarbata* in each sequence were then designated as accepted or rejected for oviposition based on the behaviour of the discriminators. One accepted and one rejected plant in each sequence was then re-tested with at least one discriminator and designated as an accepted–rejected plant pair. The accepted–rejected ranking between two plants was never reversed in re-tests (sample sizes of 1–20 discriminators were used). Natural *E. editha* attack was correlated with this preference ranking; many more naturally laid egg clusters were found on the accepted member of 25 plant pairs (30 on accepted, 9 on rejected; $\chi^2 = 11.31$, $P < 0.01$)¹¹. The rankings were also consistent between two years¹¹, and in tests of three accepted–rejected pairs with 335 females over a three-year period, all 137 discriminators agreed on the rankings.

Twenty-two accepted–rejected plant pairs were used for growing larvae. All naturally laid eggs were removed from these plants. One cluster of over 40 eggs was collected from each tested female. Upon hatching, two groups of 20 larvae were taken from each cluster and assigned to a plant pair, one group to each plant. Each of 11 plant pairs received larvae from the female that had discriminated between the two plants of that pair. The remaining 11 plant pairs each received larvae from a non-discriminator. Hatching within each cluster was synchronous, so that the two larval groups from each female were placed at the same time. Larval groups from discriminators were not paired with groups from non-discriminators, but both types were placed concurrently as they hatched during a six-day period. Natural newly hatched larvae were common at this time.

E. editha larvae feed gregariously in communal webs for ~13 days before they disperse from the web and plant. Larvae less than 13 days old are generally too small to be able to survive and feed outside the web, and are not found away from the web¹². After dispersal, the larvae feed for an additional two to three days before entering diapause for the rest of summer and winter. Because dispersing larvae cannot be followed, the number of larvae found on the plants at 10 days was used as an estimate of the immediate effect of plants on larval survival. Although survival to 10 days is not a measure of total survival or fitness, differences among plant and insect categories at 10 days may represent an important source of variation on which selection can act. The first 10 days is the period of time pre-diapause larvae are at greatest risk, with over 55% of the larval mortality overall occurring within the first five days (unpublished data and refs 11, 12).

For the offspring of discriminating females, the null hypothesis of equal survival on the two plant types is rejected (paired *t*-test; mean paired difference in the number of larvae surviving between accepted and rejected plants = 4.91, s.e.m. = 1.83, $n = 11$, $t = 2.69$, $P = 0.023$; Fig. 1a). Offspring survival of non-discriminating females was not different on the two plant types established by the discriminators (mean difference = –0.55, s.e.m. 1.36, $n = 11$, $t = -0.40$, $P = 0.70$; Fig. 1b). Overall, larval survival on each plant type was dependent on the discrimination type of the female (plant type by insect type interaction, $F = 5.76$, $P < 0.05$, Table 1).

Previous empirical studies have found that within a plant population, certain plant individuals were more likely to be attacked by the insect population, and that these plants were more suitable in general for insect offspring growth or survival^{13,14}. In addition, insect attack can decrease plant fitness^{15,16}. The interpretation is that there is selection on the insect to choose and use these more suitable plants, and selection on the

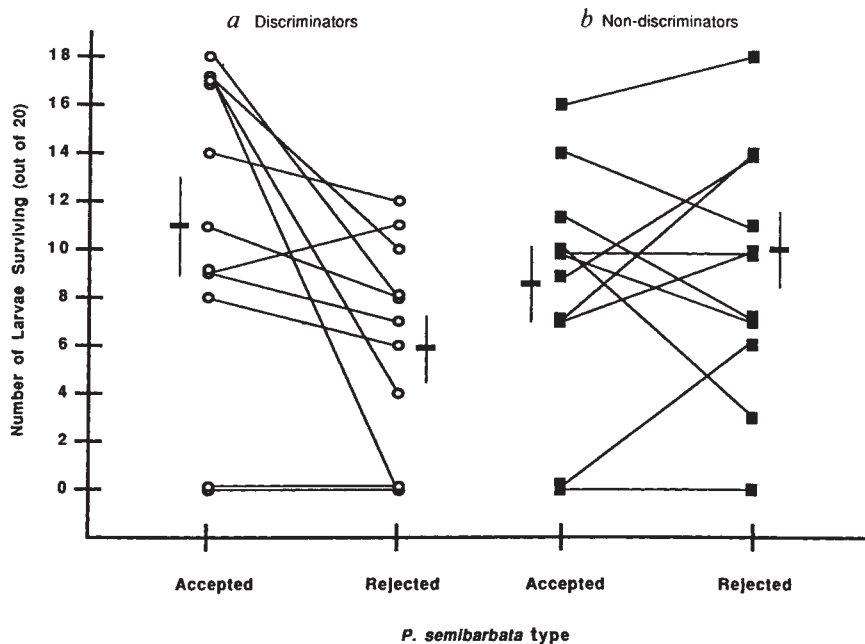


Fig. 1 Offspring survival of discriminators (a) and non-discriminators (b) split between accepted and rejected *P. semibarbata*. Lines connect full sibs. Bars to the side represent mean $\pm$ s.e.m. for larval survival of each *E. editha*-*P. semibarbata* combination.

plant to resist insect attack; though the ability to resist attack may be limited by a negative correlation between reproduction and the production of defensive compounds^{17,18}. But studies have not measured behavioural differences among insect individuals and the possible relationships between behaviour and performance on different host individuals. In this study, oviposition behaviour and offspring performance on different plant individuals was correlated, and equivalent to different host use strategies. Discriminators are specialists who preferentially lay eggs on plants that allow higher survival of their offspring. Non-discriminators are generalists who show no preference, and whose offspring survival is not correlated with oviposition choice nor with the plant categories that affect discriminators. Such specialist-generalist variation is often expected among different insect species¹⁹, but not among different individuals within a population. Yet, just as herbivorous insects need not recognize plant taxonomic categories^{13,20}, interactions between insects and plants need not follow taxonomic lines.

Failure to recognize the existence of this specialist-generalist variation within a population can result in inaccurate estimates of selection in plant-insect systems. For example, if variation among *E. editha* in oviposition behaviour were ignored, different estimates of the ability of *E. editha* to use *P. semibarbata*, or the suitability of *P. semibarbata* for *E. editha*, would have been obtained. Some *P. semibarbata* would simply have appeared more acceptable to ovipositing *E. editha* in general, and larval

survival would have been somewhat, though not significantly, higher on these more acceptable plants (mean difference = 2.18, s.e.m. = 1.26, $n = 22$, $t = 1.73$, $P = 0.10$).

In general, plants less frequently attacked by the insect population, such as the rejected *P. semibarbata* individuals, are considered to be more resistant. The presence of a generalist strategy alters this interpretation. In fact, these plants are not less acceptable to, nor less suitable for the non-discriminating generalist. Only specialist discriminators experience predictable differential survival on the two plant types. One consequence of this is that programmes for insect control that rely on increasing the proportion of 'resistant' plants may not result in the expected escape from insect attack. The proportion of generalists that can successfully attack the 'resistant' plants would simply increase. Similarly, attempts to use the mean effects of an insect population to infer evolutionary 'stalemate' or equilibrium conditions in plant defenses¹⁸ would be wrong if insects with a generalist strategy were present but not detected. These generalist insects would successfully attack and further decrease the fitness of the defended plants, disrupting any such potential equilibrium. The variation in behavioural and performance characteristics, if genetic, allows the insect to track and adapt rapidly to changes in the plant population.

I thank my dissertation committee, especially Jim Bull, John Rawlins and Mike Singer, for help and advice on the research and manuscript. D. Hinkley and P. John provided statistical advice, and N. Shah helped with fieldwork. C. Pease, C. Thomas, J. Lawton and P. Harvey also provided helpful comments. This work was supported in part by University of Texas URI and NSF grants to M. C. Singer.

Received 15 March; accepted 29 June 1988.

1. May, R. M. & Anderson, R. M. in *Coevolution* (eds Futuyma, D. J. & Slatkin, M.) 186-206 (Sinaur, Sunderland, 1983).
2. Edmunds, G. F. Jr & Alstad, D. N. *Science* **199**, 941-945 (1978).
3. Mitter, C. & Futuyma, D. J. in *Variable Plants and Herbivores in Natural and Managed Systems* (eds Denno, R. F. & McClure, M. S.) 427-459 (Academic, New York, 1983).
4. Gould, F. in *Variable Plants and Herbivores in Natural and Managed Systems* (eds Denno, R. F. & McClure, M. S.) 599-653 (Academic, New York, 1983).
5. Futuyma, D. J. & Peterson, S. A. *Rev. Ent.* **30**, 217-238 (1985).
6. Via, S. *Evolution* **40**, 778-785 (1986).
7. Singer, M. C., Ng, D. & Thomas, C. D. *Evolution* **42**, (in the press).
8. Rausher, M. D., Mackay, D. A. & Singer, M. C. *Anim. Behav.* **29**, 1220-1228 (1981).
9. Makay, D. A. *Ecol.* **66**, 142-151 (1985).
10. Singer, M. C. in *Insect-Plant Interactions* (eds Miller, J. R. & Miller, T. A.) 65-94 (Springer, New York, 1986).
11. Ng, D. Dissertation, Univ. Texas (1987).
12. Moore, S. D. Dissertation, Univ. Texas (1987).

Table 1 Analysis of variance for per cent survival of offspring

Source of variation	d.f.	SS	MS	F ratio
<i>E. editha</i> type	1	0.017	0.017	0.08
Error	20	4.385	0.219	
Plant type	1	0.155	0.155	2.74
Interaction between <i>E. editha</i> type and plant type	1	0.326	0.326	5.76*
Error	20	1.132	0.057	

A nested analysis of variance for percentage survival of offspring (arcsin $\sqrt{\text{transformed}}$) of two *E. editha* types (discriminators and non-discriminators) on two *P. semibarbata* types (accepted and rejected). Larvae from each *E. editha* female were split between the two host plant types. Total of 22 butterflies and 44 plants were used.

* $P < 0.05$.

13. Rausher, M. D. & Papaj, D. R. *Ecology* **64**, 1402-1410 (1983).
14. Myers, J. H. *J. Anim. Ecol.* **54**, 193-204 (1985).
15. Morrow, P. S. & LaMarche, V. C. *Science* **201**, 1244-1246 (1978).
16. Marquis, R. J. *Science* **226**, 537-539 (1984).
17. Futuyma, D. J. in *Coevolution* (eds Futuyma, D. J. & Slatkin, M.) 207-231 (Sinaur, Sunderland, 1983).
18. Berenbaum, M. R., Zangerl, A. R. & Nitao, J. K. *Evolution* **40**, 1215-1228 (1986).
19. Scriber, J. M. in *Variable Plants and Herbivores in Natural and Managed Systems* (eds Denno, R. F. & McClure, M. S.) 373-412 (Academic, New York, 1983).
20. Janzen, D. H. in *Herbivores: Their Interactions With Secondary Plant Metabolites* (eds Rosenthal, G. A. & Janzen, D. H.) 331-350 (Academic, London, 1979).

The variability of population densities

Stuart L. Pimm*† & Andrew Redfearn†

* Department of Zoology and † Graduate Program in Ecology, The University of Tennessee, Knoxville, Tennessee 37996, USA

The variability of population densities over time (henceforth called population variability) is one of several meanings of ecological stability¹. Here we show that estimates of the variability of population densities increase as we increase the number of years included in their calculation. This result appears for the majority of populations surveyed and over almost all the time intervals over which we calculate the variability. This result has implications for the debate over whether populations have an equilibrium.

Our analyses of population variability are motivated by a recent suggestion that, if correct, has profound implications for ecological theory: Steele's 'red-noise' metaphor². Steele demonstrates that spectral analyses of time series of physical variables in the ocean are reddened: amplitudes increase approximately as $1/(\text{frequency})$. Variation of these variables can be modelled approximately by adding a series of sine waves, with amplitudes increasing with decreasing frequency and with random phases (Fig. 1). This reddened spectrum also implies a fractal dimension to the time series³.

Now, Steele argues that, over short time scales, the spectra of terrestrial, physical variables appear white (all amplitudes are equal). Yet the time scale for the terrestrial data is much shorter than for the marine data and Williamson, for one, argues that terrestrial systems have reddened spectra too⁴. Suppose that terrestrial physical variables have reddened spectra and that physical variables force variation in population densities. Although population data are rarely long enough for sensible spectral analyses, we can ask what these suppositions mean in terms of how population variability changes over time. Irrespective of the theoretical justification, this is an important empirical question. With 'red noise' we would obtain increasing variances of population abundances as we increase the period over which the variances are calculated. What might appear as an average (an 'equilibrium') over a few years, is really a level which cycles over a longer time period, and that cycle, in turn, is moving cyclically on an even more powerful, less frequent cycle (Fig. 1).

To investigate these ideas we have determined how variability changes with the number of years included in its calculation. From the literature, we have assembled two sets of population studies that might be considered 'best' for our purposes in two different ways. The first set involves the longest studies we could find: 26 terrestrial population studies that were counted at least annually and for at least 50 years^{5,6}. There are four insect species, the rest are birds and mammals. The second set involves species counted over 24 years, all using the same technique and one designed to be the most suitable: the data on breeding densities of birds in British farmlands (42 species) and woodlands (32 species) published annually by the British Trust for Ornithology (BTO)⁷.

There are several measures of population variability⁸. We present data for the standard deviation of the logarithms of

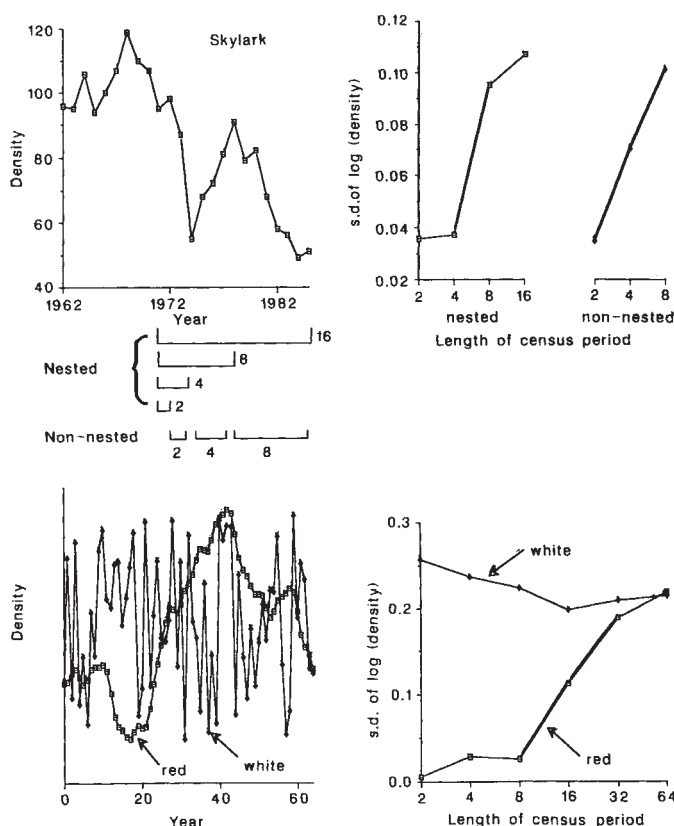


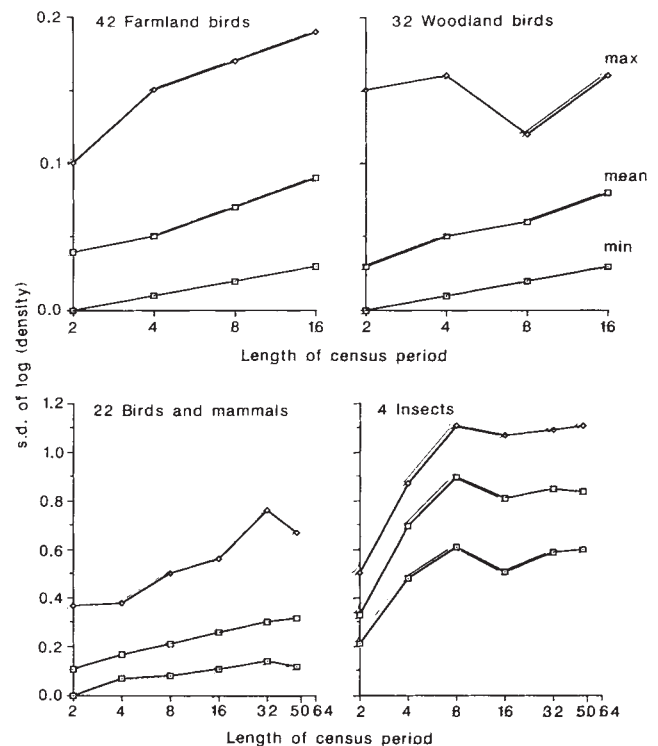
Fig. 1 Population densities and the change in variability with the length of census. *a*, Density of the skylark (*Alauda arvensis*) in English farmlands is plotted against year: the scale is relative and set to 100 in 1966. *b*, For the same population, the standard deviations of the logarithms (SDL) of density are plotted against the period over which the calculation was made. For both nested years (2, 4, 6 and 16) or non-nested years (2, 4 and 8) SDL increases with period. Pre-1970 densities were ignored because, before 1970 many bird populations were recovering from a crash in 1963 following a hard winter: inclusion of these data would have made the increase in SDL even more marked. Similar data for two model populations are shown in *c* and *d*. *c*, The two model population densities are plotted against year. The trajectory labelled 'red' was formed by summing sine waves with random phases and amplitudes that increase linearly with period. The trajectory labelled 'white' was formed by choosing densities randomly with uniform probability over the same range as those densities in the 'red' model. *d*, SDL increases for the 'red' population, but not for the 'white'.

annual density (SDL). The results for other measures are qualitatively the same.

For the data on British birds, we have calculated SDLs over 2, 4, 8 and 16 yr for each of the bird species. We have done the same calculations plus additional calculations for 32 and 50 yr for the species in the long-term data set (Fig. 2). The choice of the log (2) time scale is arbitrary. In Fig. 2, the calculations are on nested data: the first two years, the first four years, etcetera. We have also calculated SDLs for non-nested data: the first two years, the next four years, and the next eight years for the woodland and farmland birds. (We would need a total of 30 yr to be able to include estimates for SDL across the next 16 yr.) Irrespective of the method used, the population variability increases with the length of time over which we calculate it.

It is possible that this increase in average variability could be due to a small number of populations with large increases in variability. Consequently, another useful statistic is the proportion of the populations that show increases in variability with the length of the census. The majority of populations show such an increase (Table 1), so rejecting this possibility. (Only for the four species of insect is there not an increase, and this is only

Fig. 2 Population variabilities versus the number of years over which those variabilities were calculated. Only standard deviations of the logarithms of densities (SDL) are shown: the other measures of variability show a similar pattern. On each graph, the middle line joins the means among the different species; the top and bottom lines join the maximum and minimum values respectively. For the nested analyses (shown), regressions of SDL versus the length of census period are significant at $P < 0.0001$ for the farmland birds, for the woodland birds, and for the 22 miscellaneous birds and mammals; the analysis for the four insect species is significant at $P = 0.008$. For the non-nested analyses (not shown), regressions are significant at $P = 0.0001$ and $P = 0.0003$ for the farmland and woodland birds respectively. The percentages of these species that show increases between the different numbers of years that go into the calculation are shown in Table 1.



at the very longest time intervals, when the SDL levels off at about 1 (Fig. 2). It is possible that an SDL of ~ 1 is an upper limit consistent with population survival: it is equivalent to approximately four orders of magnitude of density variation over 50 yr.)

This increasing variability would be expected for populations showing simple cycles with wavelengths at least as long as the longest census period, or, comparably, if there were long-term trends in all the data. We know that some of our populations show pronounced short-term cycles. But only a small number of the species are known to cycle, and this is not a feature of any of the bird populations. In any case, the maximum cycle length is perhaps a decade (for species like the Canadian lynx), and the populations show increasing variability beyond this.

We might expect some of our populations to show long-term trends. For example, all of the populations may have been

recovering from some severe, but infrequent disturbance. The British bird populations suffered such a major collapse during the unusually cold winter of 1962/1963. To be conservative we analysed data from 1970 onwards. Alternatively, many of our populations may have been showing steady declines in abundance. Increasing variance in the bird populations, for example, might be due to degradation of say, British farmlands, with all the populations decreasing for the same reason. But there is no common trend in the species' abundances: some species increase, others decrease. Indeed, the summed densities of all the woodland and farmland species vary by less than $\pm 3\%$ in both cases. The assorted species for which we have 50 yr or more of data are from various countries in different centuries, so a single cause for the pattern of increasing variability seems unlikely.

The conclusion must be that the majority of populations show long-term trends in abundance, for different reasons and hence in different directions. Moreover, over all the time scales over which we have calculated variability, it is the long-term trends that predominate—a result strongly suggestive of these population time series having reddened spectra. Whatever the underlying model, the empirical result of increasing variance has immediate implications for such issues as how long populations persist and how long ecologically similar species may coexist.

We thank James Drake, John Gittleman, Mark Kot, John Lawton and John Steele for comments, Raymond O'Connor for supplying the BTO data, and Gregory Witteman for the figures. S.L.P. conducted this research while supported by the Institute of Nonlinear Science, University of California, San Diego. He also thanks Ted Case, Michael Gilpin and Henry Abarbanel of INLS for support, and Karin Forney and Laura Riege for making this research possible.

Table 1 Percentages of populations that increase in variability over different time spans used for the calculation

Comparison of SDLs calculated in these number of years:	2 4 8 16 32				
	versus 4	versus 8	versus 16	versus 32	versus 50
Farmland birds (42 species)	69**	72**	79**	—	—
Woodland birds (32 species)	72**	75**	72**	—	—
Various birds and mammals (22 species)	68*	68*	77***	55	41
Insects (4 species)	100	75	0	75	25

The values in the table show the percentage of the populations that increase in variability over different time spans used for the calculation. Thus, for example, the first value shows that 29 of 42 (69%) farmland bird species showed greater variabilities when calculated over 4 yr versus when calculated over just 2 yr. Over all the data we might expect the percentages to average 50%—half the populations increasing, half decreasing. On this basis, the significance of χ^2 tests are given: * $< 10\%$, ** $< 5\%$, and *** $< 1\%$; sample sizes are too small for the insect populations. The years used in the calculations are nested (see text). The measure of variability is the standard deviation of the logarithms of density (SDL).

Received 9 February; accepted 21 June 1988.

1. Pimm, S. L. *Nature* **307**, 321–326 (1984).
2. Steele, J. H. *Nature* **313**, 355–358 (1985).
3. Rothrock, D. A. & Thorndike, A. S. *J. Geophys. Res.* **85**, 3955–3963 (1980).
4. Williamson, M. H. in *Colonization, Succession and Stability* (eds Gray, A. J., Crawley, M. J. & Edwards, P. J.) 353–371 (Blackwell, Oxford, 1987).
5. Tanner, J. T. *Ecology* **47**, 733–745 (1966).
6. Peterson, R. O., Page, R. E. & Dodge, K. M. *Science* **224**, 1350–1352 (1984).
7. Marchant, J. H. & Hyde, P. A. *Bird Study* **27**, 173–178 (1980).
8. Williamson, M. H. *Ecol. Ent.* **9**, 239–241 (1984).

Inhibin β in central neural pathways involved in the control of oxytocin secretion

P. E. Sawchenko*, P. M. Plotsky†, S. W. Pfeiffer*,
E. T. Cunningham, Jr*, J. Vaughan†, J. Rivier†
& W. Vale†

* Developmental Neurobiology Laboratory and the † Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

Inhibin (I) a gonadal hormone glycoprotein which suppresses follicle-stimulating hormone (FSH) secretion from the anterior pituitary, is a heterodimer consisting of an α subunit and one of two distinct β subunits¹⁻⁵. S1 nuclease analysis has revealed that RNAs encoding all three subunits (α , β_A and β_B) are expressed in rat brain⁶. We report here on the localization, and a potential function, of inhibin β in the rat brain. A cell group centred in the nucleus of the solitary tract (NTS), a major recipient of visceral sensory information, was stained immunohistochemically with antisera against synthetic fragments of $I\beta$, but not $I\alpha$. The distribution of $I\beta$ -stained fibres is consistent with known NTS projections, and includes a prominent projection to oxytocinergic aspects of the magnocellular neurosecretory system.

The central distribution of cell bodies displaying inhibin-immunoreactivity (IR) was studied in colchicine-treated adult male and female rats. Affinity-purified antisera against synthetic $I\beta_A$ (81-113) and $I\beta_B$ (66-79) stained a small population (408 ± 29 per side; $n = 5$) of neurons centred in the commissural part of the NTS, near the spinal-medullary junction (Fig. 1). At similar levels, a smaller complement (143 ± 18) of $I\beta_A$ -stained

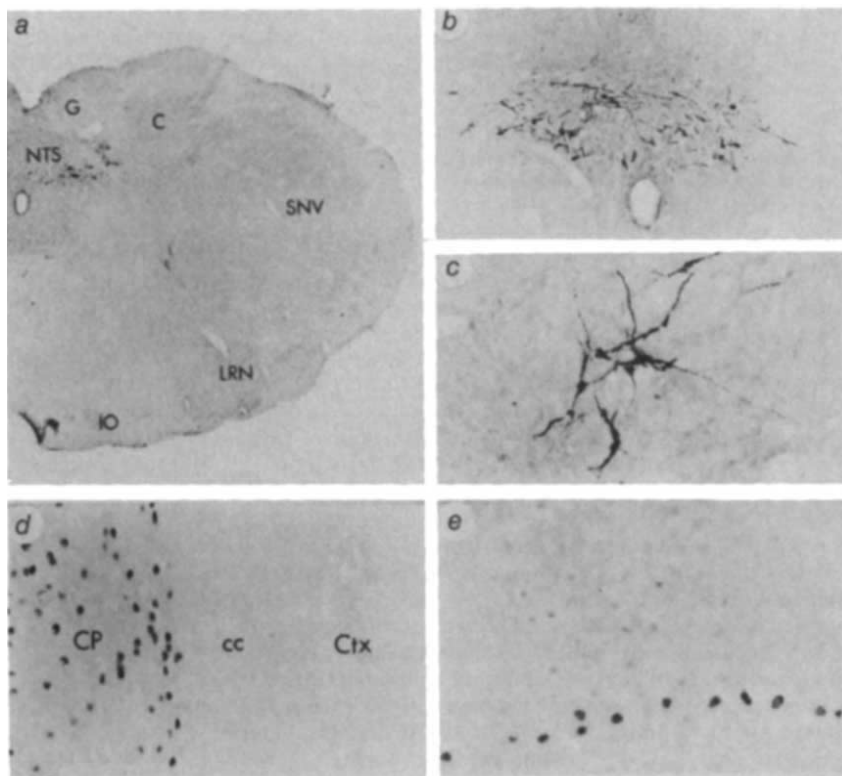
cells was found in the medullary reticular formation. These were the only cells in the brain and spinal cord that displayed specific perikaryal and dendritic staining using conventional immunofluorescence or biotin-avidin immunoperoxidase methods⁷. Widespread staining of neuronal nuclei was also evident. Some regional specificity in the distribution of nuclear staining was apparent, with the caudoputamen and cerebellum comprising two foci; the well known axonal projections of these areas did not display $I\beta_A$ -IR. Nuclear localization of an agent which is a secreted protein would be unusual, though not unprecedented⁸.

Axons and terminals displaying $I\beta_A$ -IR were visualized in untreated rats (Fig. 2), although, compared to most neuroactive substances, they were not numerous. Their distribution was consistent with known efferent projections of the NTS and/or the medullary reticular formation⁹⁻¹¹. Because inhibin-related peptides are associated with reproductive function, it is interesting that $I\beta$ -stained terminals were detected in those aspects of the septal-preoptic region where neurons containing gonadotropin-releasing hormone are found¹², and, more prominently, in those regions of the paraventricular (PVN) and supraoptic nuclei (SON) where oxytocin-containing magnocellular neurosecretory neurons are concentrated (Fig. 2).

Because molecules as large as inhibin or activin dimers are not usually synaptically active, conventional pre-embedding immunoperoxidase methods were used to localize $I\beta$ -IR ultrastructurally in the PVN and SON (Fig. 3). In both nuclei, numerous presynaptic profiles containing predominantly small (25-60 nm), round, electron-lucent vesicles and a few (0-5) larger (70-100 nm) dense-cored vesicles stained positively for $I\beta$ -IR. Labelled terminals were of both the symmetric and asymmetric types, and most ended on dendritic shafts. Examples of axosomatic contacts, some with magnocellular neurosecretory somata, were encountered. Dual immunoperoxidase and immuno-autoradiographic localization of oxytocin- and $I\beta$ -IR,

Fig. 1 Brightfield photomicrographs of $I\beta$ staining of neuronal cell bodies in rat brain. *a*, Hemisection ($\times 18.6$) of the caudal medulla showing stained perikarya in the NTS and reticular formation in a colchicine-treated rat. These were the only sites of cytoplasmic and dendritic staining detected in brain. C, cuneate nucleus; G, gracile nucleus; IO, inferior olive; LRN, lateral reticular nucleus; SNV, spinal trigeminal nucleus. *b*, *c*, Higher-power views of neurons in the NTS (*b*, $\times 46.5$) and reticular formation (*c*, $\times 116.5$). *d*, *e*, Neuronal nuclei displaying $I\beta$ -IR in the caudoputamen (CP, *d*) and cerebellum (*e*). Though apparent in many brain regions, nuclear staining showed some regional specificity. Many stained nuclei were evident in the striatum (left side of *d*), but not in adjacent deep layers of cerebral cortex (Ctx, right side). cc, Corpus callosum. In the cerebellum, most Purkinje cell nuclei (bottom of *e*) and some nuclei in the molecular layer (top) were stained ($\times 186$).

Methods. Animals received four ventricular injections of colchicine (50 μ g in 25 μ l saline) 48-72 h before they were killed. They were then anaesthetized and perfused transcardially with 500 ml ice-cold 4% paraformaldehyde in borate buffer at pH 9.5. After overnight postfixation, brains were frozen and sectioned at 30 μ m. Immunostaining was carried out using primary antisera against $I\beta_A$ (81-113) raised in rabbits diluted at 1:2,000. A biotin-avidin-immunoperoxidase method⁷ (Vector) was used for localization. Both cytoplasmic and nuclear localizations were abolished by preadsorption of the antisera with 60 μ M of their respective synthetic immunogens, with equimolar concentrations of purified¹³ $I\beta_A$ homodimer, and the corresponding synthetic sequence of $I\beta_B$ (80-112). The minimum concentrations for competition of staining with $I\beta_A$ (81-113) and $I\beta_B$ (80-112) were 0.3 μ M and 30 μ M, respectively. Pre-adsorption of the antiserum with 60- μ M concentrations of structurally related human TGF- β^2 (R&D Systems, Minneapolis) failed to interfere with staining. Antisera against porcine inhibin α (1-26), which stain granulosa cells in rat or porcine ovary strongly (data not shown), consistently failed to label neurons in the medulla. The anti- $I\beta$ sera used here also stain ovarian granulosa cells, though less intensely than sera against $I\alpha$.



Antisera against porcine inhibin α (1-26), which stain granulosa cells in rat or porcine ovary strongly (data not shown), consistently failed to label neurons in the medulla. The anti- $I\beta$ sera used here also stain ovarian granulosa cells, though less intensely than sera against $I\alpha$.

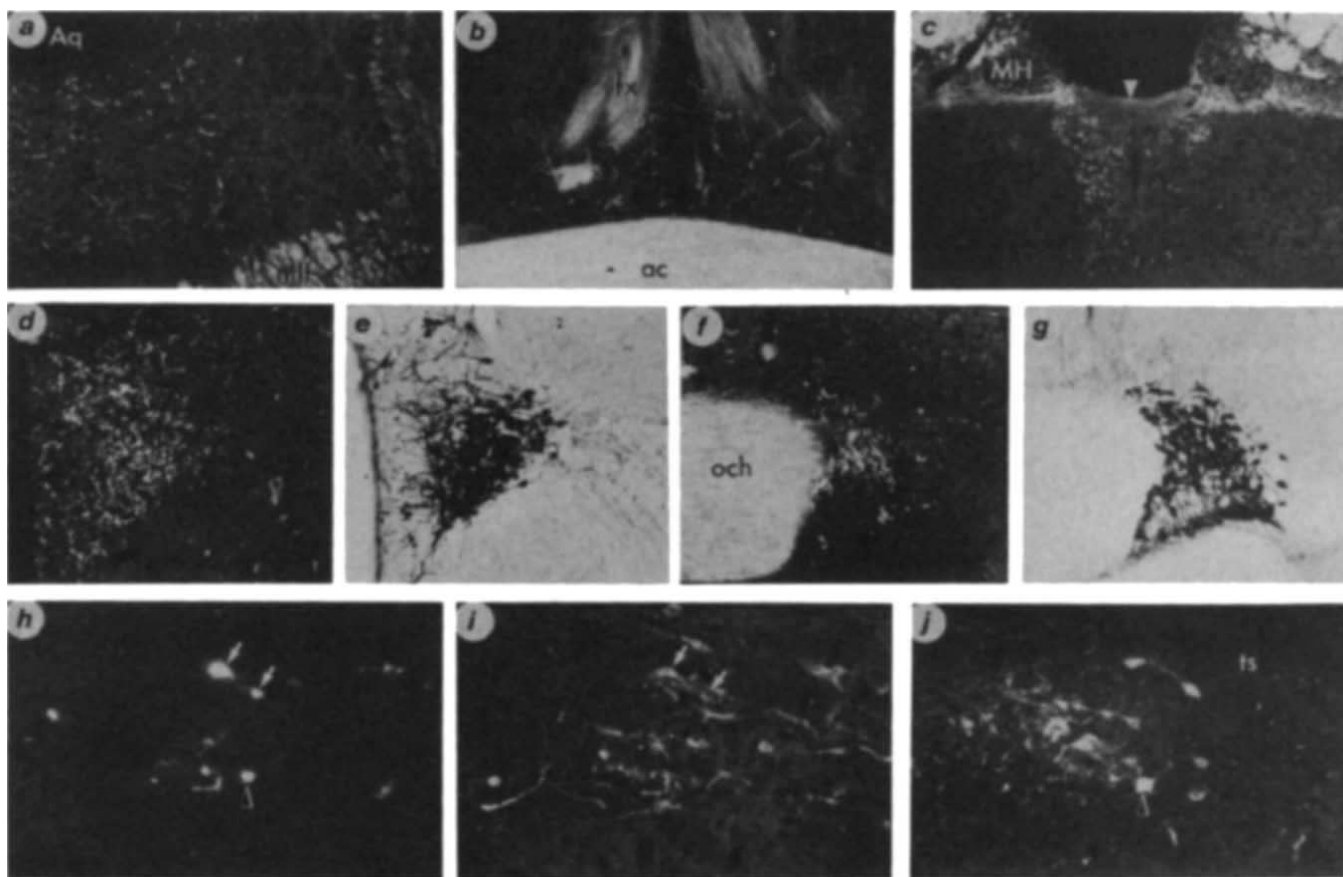


Fig. 2 *a-c*, Darkfield photomicrographs of $I\beta$ -stained fibres and terminals in rat brain. *a*, Pontine central grey at the level of the dorsal raphe nucleus. Aq, cerebral aqueduct; mlf, Medial longitudinal fasciculus. *b*, Median preoptic nucleus. ac, anterior commissure; fx, fornix. *c*, Paraventricular nucleus of the thalamus at the level of the medial habenula (MH) in a rat that received a unilateral brainstem knife cut. Note depletion of fibres on the lesioned (right) side. Arrowhead indicates midline ($\times 42$). *d-g*, Relationship of $I\beta_A$ -stained terminal fields in the PVN and SON to magnocellular oxytocin neurons. *d, e*, Adjacent sections through the PVN to show $I\beta$ -IR fibres (left) and oxytocin-IR perikarya (right). A prominent input is evident. *f, g*, Similar comparison in the SON. Oxytocin neurons are concentrated in the dorsal aspect of the nucleus, as are $I\beta$ -IR terminals ($\times 65$). *h-j*, Fluorescence photomicrographs of a single section through the NTS viewed under different illumination conditions to show cells retrogradely labelled after tracer deposits in the PVN (*h*), and cells stained for $I\beta$ (*i*) and DBH (*j*). Arrows indicate retrogradely labelled cells displaying $I\beta_A$ -IR; the arrowhead indicates a retrogradely labelled cell stained for DBH. No co-localization of $I\beta$ - and DBH-IR is evident ($\times 83$).

Methods. Avidin-biotin-immunoperoxidase staining carried out as described in Fig. 1. Combined retrograde transport-immunofluorescence technique described elsewhere¹⁸. Knife cuts²⁸ were placed in the coronal plane at the level of the rostral pole of the facial nucleus extending 2 mm medially from the medial aspect of the spinal trigeminal nucleus; depth of cut was 2 mm. Animals were allowed 14–17 days' survival before perfusion. Oxytocin staining was as previously described. Staining was abolished by preincubation with $10 \mu\text{g ml}^{-1}$ synthetic oxytocin, but not vasopressin.

respectively, showed that $I\beta$ -IR terminals do terminate upon oxytocin-IR elements in both magnocellular nuclei (Fig. 3). The selectivity of this innervation remains to be established.

All localizations were abolished by pre-adsorption of antisera with their synthetic immunogens and with activin A ($I\beta_A$ homodimer) purified from porcine follicular fluid¹³. Adsorption of anti- $I\beta_A$ (81–113) with the corresponding synthetic $I\beta_B$ sequence (80–112) interfered with staining, but 100-fold less effectively than synthetic $I\beta_A$ (81–113); we nevertheless refer to the agent(s) recognized by our antisera generically as $I\beta$. Apart from a pronounced structural similarity to transforming growth factor- β ¹⁴ (TGF- β), inhibin subunits show sequence homologies with mullerian inhibitory substance¹⁵ and the decapentaplegic gene complex of *Drosophila*¹⁶. TGF- β failed to interfere with staining, indicating that the material recognized by our antisera is $I\beta$ or an unidentified and highly related congener. The failure of antisera against porcine $I\alpha$ to provide positive staining suggests that the inhibin-like peptide recognized by our sera is activin, which exists as β -homodimers^{13,17}, rather than inhibin.

Catecholaminergic neurons are the most thoroughly characterized^{14,18} output neurons of the NTS, so we used a combined retrograde transport-double immunohistochemical staining technique to compare the number and distribution of cells immunoreactive for $I\beta$ and dopamine- β -hydroxylase (DBH) in the medulla that could be labelled after deposit of a retrogradely transported fluorochrome in the PVN (Fig. 2). Some 9% of the total number of retrogradely labelled cells in the NTS (44 ± 12 ; $n = 6$) and 3% of those in the reticular formation (16 ± 6) were immunoreactive for $I\beta$, but not DBH. There was no evidence, in any of the 6 separate experiments, for co-localization of DBH- and $I\beta$ -IR. These results indicate that $I\beta$ -IR neurons in the caudal medulla are distinct from catecholaminergic cells, and contribute to at least one prominent terminal field in the forebrain. Unilateral medullary knife cuts, which proved successful in depleting DBH-IR from the PVN, were also found to greatly reduce $I\beta$ -IR in the forebrain (Fig. 2), supporting the hypothesis. Parvocellular neurosecretory neurons and vasopressin-containing magnocellular neurosecretory neurons are prominent targets for ascending catecholaminergic projections to the

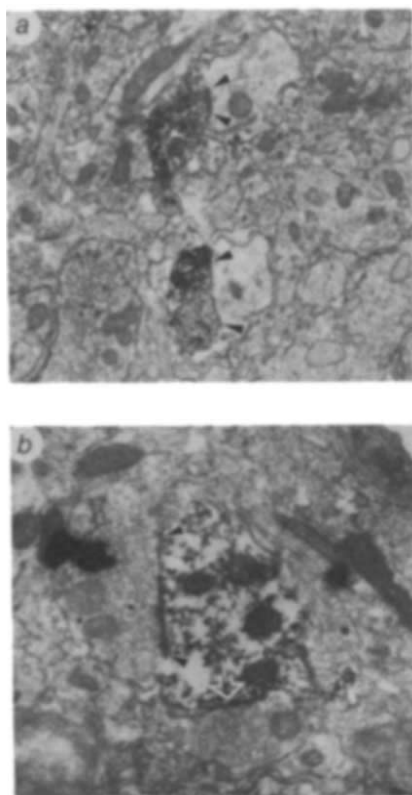


Fig. 3 Ultrastructural localization of $I\beta$ -IR in terminals in the PVN and SON. *a*, Pre-embedding immunoperoxidase staining of terminals in the SON. Profiles of two labelled terminals in synaptic contact (arrowheads) with unlabelled dendrites ($\times 12,600$). *b*, Autoradiographic localization of $I\beta$ -IR in a terminal that makes a synaptic contact (arrowheads) with a dendrite in the PVN which is stained positively (using immunoperoxidase methods) for oxytocin-IR. Reduced silver grains (irregular and dark) mark the $I\beta$ -IR terminal; an unlabelled terminal contacting the same dendrite is indicated (open arrow) ($\times 15,400$).

Methods. Tissue from five adult male rats was prepared for single and dual ultrastructural labelling using a protocol described by Pickel and colleagues²⁹. Single (immunoperoxidase) labelling was carried out on 50- μ m-thick vibratome sections using the same anti- $I\beta_A$ serum used for light-level studies. For double labelling studies, rabbit anti- $I\beta$ and swine anti-oxytocin (Accurate Chemicals) were applied to tissue concurrently, and localized using an 125 I-labelled donkey anti-rabbit immunoglobulin G (IgG) ($10 \mu\text{Ci } \mu\text{g}^{-1}$; Amersham) and biotinylated goat anti-swine IgG (Vector Labs), respectively. Ultrathin sections were collected on grids, coated with emulsion (Ilford L4), and exposed for 4–6 weeks. All sera used in ultrastructural studies met the criteria for specificity described above for light-level localizations.

PVN^{11,18–20}. In suggesting that $I\beta$ -IR cells in the caudal NTS project to oxytocin neurons, the results provide further evidence for chemically differentiated pathways by which peripheral stimuli may evoke specific endocrine responses.

The results suggest that an $I\beta$ -like peptide is present in the rat brain and is able to participate in both intra- and inter-neuronal functions. We have concentrated on the latter, describing a distribution of $I\beta$ -IR perikarya and fibres which suggests a role in specific aspects of sensory information processing and showed that this pathway interacts synaptically with oxytocinergic neurons. The pattern of $I\beta$ -IR that we have described is unique; no other chemically defined neural system outside the hypothalamus is present in cell bodies with so limited a distribution. It is unknown whether this or other inhibin-related peptides are expressed at levels beneath the detection limits of our

methods, or are induced by manipulation of physiological status. Inhibin α RNAs are more abundant in whole rat brain than either β species⁶ indicating that our knowledge of the central distribution of inhibin-related peptides is not complete.

The modalities of sensory information that may be conveyed to $I\beta$ -IR neurons in the NTS remain to be specified. No single visceral afferent is known to terminate discretely in the region occupied by $I\beta$ -IR neurons, though the sensory distribution of the vagus encompasses this region²¹. It is uncertain whether somatic sensory information transmitted initially to the spinal cord may reach $I\beta$ -IR cells, though pathways from the dorsal horn to the caudal NTS have been described²².

The most thoroughly characterized functional roles of oxytocin are in the neuroendocrine control of milk ejection and parturition. Despite advances in charting potential pathways by which somatosensory input from the nipples, for example, may be conveyed to oxytocinergic neurons^{23,24}, no substantial ascending inputs to oxytocin-containing cell groups have been identified previously. Preliminary results indicate that infusion of femtomolar quantities of purified activin, but not inhibin, into the PVN can elicit oxytocin secretion in anaesthetized male rats, and that local infusion of anti- $I\beta$ sera can attenuate suckling-induced oxytocin secretion in anaesthetized dams²⁵. With the anatomical evidence presented here, these results justify further examination of a role for activin in suckling-induced oxytocin secretion. This pathway need not necessarily comprise the central neural limb of the milk ejection reflex, and other functional associations are possible. For example, gastrointestinal stimuli related to ingestion or nausea have been shown to stimulate oxytocin secretion²⁶ and firing of identified oxytocinergic neurons²⁷. These effects are vagally mediated²⁶, and easily accommodated by the circuitry described here. Potential associations of inhibin β -IR pathways with other visceral regulatory functions remain to be investigated.

This work was supported by the NIH, by a McKnight Foundation Scholar's Award, by a contract from the US-AID and the Population Council, and was conducted in part by the Clayton Foundation for Research, California Division. P.E.S., P.M.P. and W.V. are Clayton Foundation Investigators. We thank Dr R. Grzanna for providing antisera against DBH, and acknowledge the expert technical assistance of Carlos Arias.

Received 19 February; accepted 30 June 1988.

- Mason, A. J. *et al.* *Nature* **318**, 659–663 (1985).
- Mason, A. J., Niall, H. D. & Seeburg, P. H. *Biochem. biophys. Res. Commun.* **135**, 957–964 (1986).
- Mayo, K. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5849–5853 (1986).
- Forage, R. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 3091–3095 (1986).
- Esch, F. S. *et al.* *Molec. Endocr.* **1**, 388–396 (1987).
- Meunier, H. F., Rivier, C., Evans, R. M. & Vale, W. *Proc. natn. Acad. Sci. U.S.A.* **85**, 247–251 (1988).
- Gerfen, C. R. & Sawchenko, P. E. *Brain Res.* **290**, 219–238 (1984).
- Yeh, H.-J., Pierce, G. F. & Deuel, T. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2317–2321 (1987).
- Ricardo, J. & Koh, E. T. *Brain Res.* **153**, 1–26 (1978).
- McKellar, S. & Loewy, A. D. *Brain Res.* **241**, 11–29 (1982).
- Sawchenko, P. E. & Swanson, L. W. *Brain Res. Rev.* **4**, 275–325 (1982).
- Witkin, J. W., Paden, C. M. & Silverman, A. J. *Neuroendocrinology* **35**, 429–438 (1982).
- Vale, W. *et al.* *Nature* **321**, 776–779 (1986).
- Derynck, R. *et al.* *Nature* **316**, 701–785 (1985).
- Cate, R. L. *et al.* *Cell* **45**, 685–698 (1986).
- Padgett, R. W., St. Johnston, R. D. & Gelbart, W. M. *Nature* **325**, 81–84 (1987).
- Ling, N. *et al.* *Nature* **321**, 779–782 (1986).
- Cunningham Jr., E. T. & Sawchenko, P. E. *J. comp. Neurol.* (in the press).
- Swanson, L. W. *et al.* *J. comp. Neurol.* **196**, 271–285 (1981).
- Liposits, Zs., Phelix, C. & Paull, W. K. *Histochemistry* **84**, 105–120 (1986).
- Contreras, R. J., Beckstead, R. M. & Norgren, R. J. *Auton. Nerv. Syst.* **6**, 302–322 (1982).
- Menetrey, D. & Basbaum, A. I. *J. comp. Neurol.* **255**, 438–450 (1987).
- Dubois-Dauphin, M., Armstrong, W. E., Tribollet, E. & Dreifuss, J. J. *Neuroscience* **15**, 1111–1129 (1985).
- Tasker, J. G., Theodosis, D. T. & Poulain, D. A. *Neuroscience* **19**, 495–509 (1986).
- Plotsky, P. M., Sawchenko, P. E. & Vale, W. *Soc. Neurosci. Abstr.* (in the press).
- Verbalis, J. G., McCann, M. J., McHale, C. M. & Stricker, E. M. *Science* **232**, 1417–1419 (1986).
- Renaud, L. P., Tang, M., McCann, M. J., Stricker, E. M. & Verbalis, J. G. *Am. J. Physiol.* **253**, R661–R665 (1987).
- Gold, R. M., Kapatos, G. & Carey, R. J. *Physiol. Behav.* **10**, 813–815 (1973).
- Pickel, V. M., Chan, J. & Milner, T. A. *J. Histochem. Cytochem.* **34**, 792–806 (1986).

Protein kinase C mediates neural induction in *Xenopus laevis*

A. P. Otte, C. H. Koster, G. T. Snoek & A. J. Durston

Hubrecht Laboratory, Netherlands Institute for Developmental Biology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Inductive cell interactions are essential in early embryonic development, but virtually nothing is known about the molecular mechanisms involved^{1,2}. Recently factors resembling fibroblast growth factor³⁻⁶ and transforming growth factor- β ⁴⁻⁶ were shown to be involved in mesoderm induction in *Xenopus laevis*, suggesting that membrane receptor-mediated signal transduction is important in induction processes. Here we report direct measurements of protein kinase C (PKC) activity in uninduced ectoderm, and in neuroectoderm shortly after induction by the involuting mesoderm, in *Xenopus laevis* embryos. Membrane-bound PKC activity increased three to fourfold in the induced neuroectoderm while the cytosolic PKC activity was decreasing, indicating that PKC activity was translocated during neural induction. A similar time- and dose-dependent translocation of activity was seen after incubation with the PKC activator 12-*O*-tetradecanoyl phorbol-13-acetate⁷, which also induced neural tissue in competent ectoderm, suggesting that PKC is involved in the response to the endogenous inducing signal during neural induction.

Accumulating evidence indicates that cell-surface mediated signal transduction²⁻⁶ is involved in induction processes in early embryogenesis. Several growth-factors and hormones that stimulate inositol phospholipid turnover after receptor activation induce translocation of PKC from the cytosol to the membrane, leading to activation of PKC (ref. 8). If similar, but as yet unidentified, growth-factor receptors are involved in embryonic induction processes, then they should be associated with PKC translocation. We tested this hypothesis for neural induction in *Xenopus laevis*.

An *in vitro* assay for *Xenopus laevis* Ca^{2+} - and phospholipid-dependent PKC activity was developed. PKC activity was measured in the particulate (membrane) and soluble (cytosolic) fractions^{9,10} from whole embryos, dissected uninduced ectoderm and induced neuroectoderm, and *in vitro* induced neuroectoderm. In all cases, two developmental stages that mark the beginning and end of neural induction were compared: early gastrula (stage 10) and early neurula (stage 13) (ref. 11).

As shown in Table 1 and Fig. 1a, the period of neural induction (stage 10 to stage 13) is characterized by a threefold increase in total and specific membrane-bound PKC activity and an almost twofold decrease in cytosolic total and specific PKC activity, as determined in whole embryo preparations. This enzyme may thus be implicated in the transduction of the underlying inducing signal.

During gastrulation the involuting mesoderm induces ectoderm to become neural tissue. We took advantage of the fact that it is still possible to remove the induced neuroectoderm at the end of gastrulation (stage 13) free of contamination by the underlying mesoderm. We measured PKC activity in uninduced stage 10 ectoderm and in dissected induced stage 13 neuroectoderm. We found an increase in translocation of PKC activity compared with whole embryos (Fig. 1b). We also measured PKC activity in the remaining part of the embryo consisting of endoderm and mesoderm, and ventral ectoderm as well in stage 13. We found no significant changes in cytosolic (S100, stage 10: 50 ± 10 pmol min⁻¹ mg protein⁻¹, $n = 5$; stage 13: 50 ± 10 pmol min⁻¹ mg protein⁻¹, $n = 5$) or in membrane-bound (P100, stage 10: 60 ± 10 pmol min⁻¹ mg protein⁻¹, $n = 5$; stage 13: 80 ± 15 pmol min⁻¹ mg protein⁻¹, $n = 5$) PKC activities. This indicates that PKC translocation is associated with neural induction and not with other events in the non-neural remainder

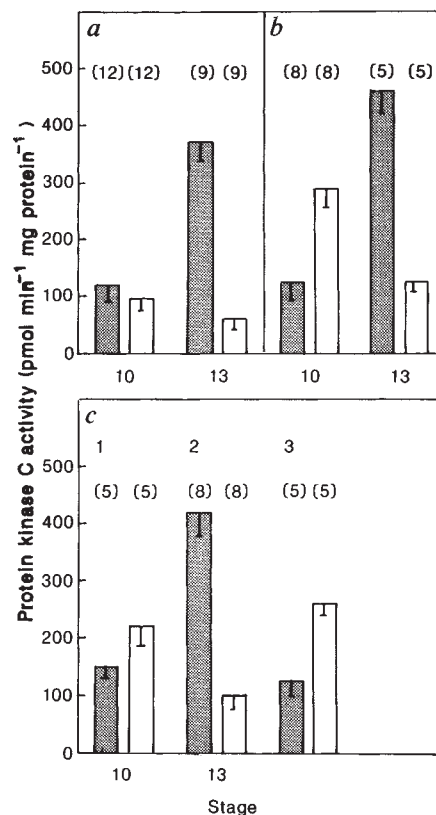


Fig. 1 Translocation of PKC activity from the cytosol to the membrane fraction of *Xenopus laevis* embryos during neural induction. PKC activity was measured in membrane (cross-hatched) and cytosolic fractions (open) from a, whole embryos at stage 10 and 13 (ref. 11); b, ectoderm from stage 10 and neuroectoderm from stage 13; c, ectoderm from stage 10 with a small piece of dorsal marginal zone left attached (1 and 2). PKC activities were measured either immediately (1), or after the explants were cultured long enough for control stage 10 embryos to reach stage 13 (2). As a negative control, ectoderm explants without marginal zone mesoderm were cultured for the same time before measuring PKC activity (3). PKC activity was defined as Ca^{2+} - and phospholipid-dependent kinase activity and is expressed as pmol ^{32}P incorporated into the substrate (histone III S) per min per mg protein in the enzyme preparation. Values are means $\pm$ standard errors of n independent experiments (shown in parentheses) using 40 embryos or explants per measurement. All reactions contained 5 μg protein per assay and were carried out in duplicate for 5 and 15 min. **Methods.** Embryos or explants were washed and homogenized in ice-cold lysis buffer^{9,10}. The homogenate was centrifuged at 10,000g for 15 s, the pellet homogenized again, centrifuged and the supernatants combined. The pellet was discarded as it contained no PKC activity. The supernatant was spun at 100,000g for 1 hour. The 100,000g supernatant, considered to be the cytosolic fraction, was removed, and the pellet resuspended and slowly rotated for 30 min in lysis buffer containing 1% Nonidet-40. The mixture was centrifuged for 15 min at 10,000g and the pellet discarded. The supernatant was considered to be the particulate or membrane fraction. Partial purification of PKC with DEAE-cellulose and the subsequent PKC activity assay were performed as described^{9,10}. Elution was with lysis buffer containing 70 mM NaCl with which >90% of both soluble and particulate PKC activity eluted. We routinely determined the ratio between total protein content of the material loaded onto the DEAE column and the protein content of the eluted DEAE fraction. Between the two developmental stages that we compared, we found these ratios to be the same for both stages, both for the particulate and cytosolic fractions. These ratios were also constant for particulate and cytosol fractions respectively from explants of all types and ages and dissected ectoderm. This means that the changes in activity in the DEAE fraction reflect activity changes in the total protein sample before DEAE fractionation. Further, as the protein contents of whole embryos and of explants did not change between stage 10 and 13, the time-dependent changes that we report as specific activities in the DEAE fractions accurately reflect changes in total activities. We also present total PKC activities for some specific cases in Table 1. The enzyme activity was linear with time up to 60 min.

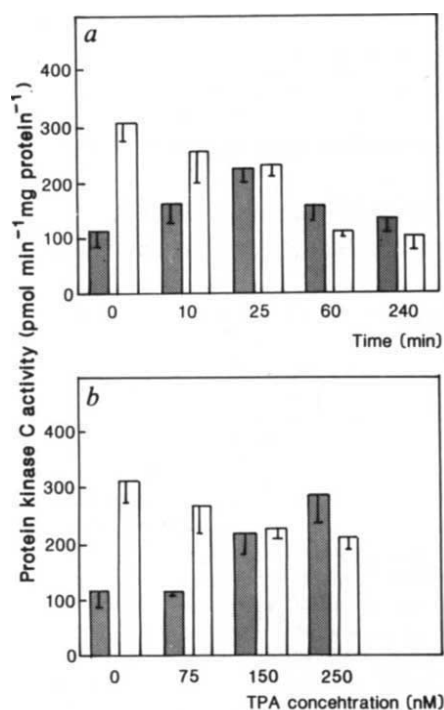


Fig. 2 Effect of TPA on PKC activities in stage 10 ectoderm. Explants of stage 10 ectoderm were treated with 150 nM TPA in 100% Flickinger solution²³ for the time indicated (a) or for 25 minutes with the concentrations indicated (b). After the incubation, explants were washed with Flickinger solution, homogenized immediately and treated as described in Fig. 1 legend. Values are means $\pm$ standard errors of at least 4 independent experiments.

of the embryo during gastrulation. Next we tested whether the translocation of PKC activity was a direct result of neural induction, or simply a consequence of the ageing of the ectoderm. Stage 10 ectoderm explants were cultured for ~5 h, which was long enough for complete control stage 10 embryos to reach stage 13. Measurement of PKC activities showed that no translocation had occurred (Fig. 1c, 3), ruling out the possibility that any translocation observed resulted from ageing of the ectoderm. We also dissected the same amount of stage 10 ectoderm with a small piece (less than 20% of the explant) of dorsal marginal zone mesoderm attached. Culturing with dorsal marginal zone mesoderm gives neural tissue, as shown by histology (result not shown), and PKC activities measured in such ectomesoderm explants directly after dissection and after culturing for 5 h showed extensive translocation to the membrane. (Table 1 and Fig. 1c, 1 and 2), indicating that the translocation results directly from the neural inducing action of the mesoderm.

We have also compared total PKC activities in responsive ectoderm before and after neural induction, the initial amount of tissue being the same in each case. Table 1 shows that the total membrane-associated PKC activity increases by more than

Table 1 Total PKC activities during neural induction

Stage	P100	S100	P+S	% P/(P+S)
a, Whole embryo 10	4 $\pm$ 0.8	20.5 $\pm$ 3	24.5	16
Whole embryo 13	12 $\pm$ 2	16 $\pm$ 1.5	28	43
b, Ect + mz (10)	2.5 $\pm$ 0.5	11 $\pm$ 2	13.5	18
Ect + mz (13)	9 $\pm$ 1.5	7 $\pm$ 1.5	16	56

Total activities were measured in particulate (P100) and cytosolic (S100) DEAE fractions of a, whole embryos of stage 10 and 13 and b, ectoderm plus dorsal marginal zone, either measured immediately after dissection (10) or after culturing till control embryos reached stage 13 (13). Activities are expressed as pmol ³²P incorporated into substrate per min, in the total DEAE fraction of 100 embryos or explants (mean $\pm$ s.e., *n* experiments, as in Fig. 1a and c).

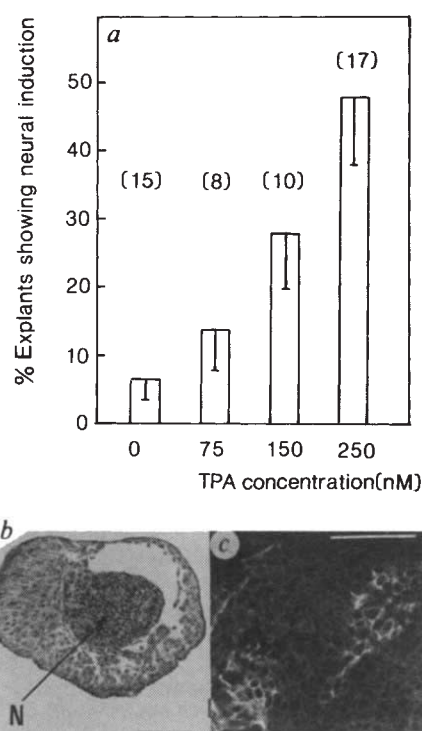


Fig. 3 Induction of neural tissue by TPA in *Xenopus laevis* ectoderm. a, Explants of stage 10 ectoderm were incubated with the TPA concentrations indicated, as described in Fig. 2, for 4 or 16 h at 20 °C. The percentage of explants showing neural induction are shown. No substantial differences were found between the two incubation times. Values are means $\pm$ standard errors of *n* independent experiments (shown in parentheses) using 10–20 explants per experiment. b, Histological section of a treated explant induced with 250 nM TPA showing neural tissue (N). Following TPA treatment, the explants were cultured for two days at 20 °C with four changes in Flickinger medium before fixation in Smith's fixative (2.5% acetic acid, 0.5% K₂Cr₂O₇, 4% formaldehyde) and haematoxylin-eosin staining. We find no toxicity using overnight exposure of stage 10 ectoderm explants to TPA up to a concentration of 500 nM. No dead cells are seen in TPA-treated or control explants after overnight culture as deduced from uptake of trypan blue or propidium iodide, and no necrotic cells or cell debris are found after 2 days in culture as judged by haematoxylin-eosin staining and histology. Scale bar: 100 μ m. c, Immunostaining of neural-induced tissue by 250 nM TPA with monoclonal antibody 2G9 which reacts specifically with neural tissue, fixation was in 2% trichloroacetic acid. Scale bar: 50 μ m.

threefold in neural induced ectoderm and that this increase represents a large fraction of the translocated activity in the whole embryo. This emphasizes the correlation between neural induction and translocation of PKC activity to the membrane fraction. No changes in total membrane-associated PKC activity were observed in control (uninduced, marginal zoneless) ectoderm over the same time period (results not shown).

If the translocation of PKC activity is causally related to neural induction, it is possible that drug-induced translocation of PKC activity could cause neural induction. To test this hypothesis, we incubated uninduced stage 10 ectoderm with 12-*O*-tetradecanoylphorbol-13-acetate (TPA), a potent activator of protein kinase C, which is known to induce translocation of the enzyme to the plasma membrane in mammalian cells^{12,13}. We found that the induced translocation of PKC activity from the cytosol to the membrane fraction by TPA was time-dependent (Fig. 2a), reaching a maximum after 25 min, the membrane and cytosolic PKC activity both decreasing at longer times. The effect of TPA was dose-dependent: at 75 nM, TPA did not increase the membrane-associated PKC activity, but after 25 min incubation with 250 nM TPA (the highest concentration tested),

the membrane-bound PKC activity more than doubled (Fig. 2b). These data correlate with the ability of TPA to cause neural induction. 250 nM TPA induced neural tissue in 50% of the treated explants (Fig. 3a), whereas the biologically inactive phorbol ester 4- α -phorbol-12,13-didecanoate had no effect at 20 μ M. The presence of neural tissue was demonstrated by classical histology (Fig. 3b) and with a monoclonal antibody that reacts specifically with neural tissue (Fig. 3c). The number of induced explants and the amount of neural tissue per explant were both dose-dependent. In agreement with the observed lack of any PKC translocation after 25 min, it was found that 75 nM TPA caused barely any neural induction in explants. These results differ in detail from the recently described effects of TPA on neural induction in ectoderm of another amphibian species, *Triturus alpestris*, which seems to be more sensitive than *Xenopus laevis* to the action of TPA (ref. 14).

We effectively conclude that mesoderm, the natural neural inducer, induces translocation of PKC activity from the cytosolic to the particulate fraction. TPA is known to induce translocation and activation of PKC in mammalian cells, and we find that it also induces translocation of *Xenopus laevis* PKC and causes neural induction. Evidently, PKC translocation is caused for the induction of neural differentiation. We conclude that activated growth factor-like receptors that cause PKC-activation are involved in embryonic induction.

Mammalian cell lines typically show relatively rapid transient PKC translocation on activation⁸, whereas *Xenopus* neural induction is associated with slow, but relatively persistent, PKC translocation. Only one report demonstrates similar PKC translocation kinetics: after induction of long-term potentiation, a process mimicked by TPA (ref. 15), there was a persistent translocation of PKC activity to membranes prepared from whole brain tissue one hour after the stimulus¹⁶.

To our knowledge, this is the first report of PKC translocation *in vivo* resulting from the transduction of an endogenous signal during the regulation of a naturally occurring process and it further emphasizes the importance of signal transduction in embryogenesis³⁻⁶. Furthermore, several genes that are important in embryonic pattern formation, such as DPP-C (ref. 17), *sevenless*¹⁸ and *Notch*¹⁹, show homology with genes coding for growth factors or their receptors. Proto-oncogenes, like *int-1* and *int-2*, that might code for growth factors²⁰ also seem to be important in embryonic development^{21,22}. Investigation of signal transduction pathways is a promising new approach to the clarification of the mechanisms underlying embryonic induction and pattern formation.

We thank Dr E. A. Jones for the monoclonal antibody 2G9; W. van Rotterdam and A. Wicherink for help in developing the PKC assay, P. van Run for Fig. 3c, and Drs R. van Driel, P. Nieuwkoop, W. Moolenaar, S. de Laat, M. Heideveld and P. van der Saag for discussion.

Received 19 May; accepted 15 June 1988.

1. Nieuwkoop, P. D. *Curr. Topics dev. Biol.* **11**, 115-32 (1977).
2. Gurdon, J. B. *Development* **99**, 285-306 (1987).
3. Slack, J. M. W., Darlington, B. G., Heath, J. K. & Godsave, S. F. *Nature* **326**, 197-200 (1987).
4. Weeks, D. L. & Melton, O. A. *Cell* **51**, 861-867 (1987).
5. Kimelman, D. & Kirschner, M. *Cell* **51**, 869-877 (1987).
6. Rosa, F. *et al. Science* **239**, 783-785 (1988).
7. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
8. Niedel, J. E. & Blackshear, P. J. in *Phosphoinositides and Receptor Mechanisms* (ed. Putney, J. W.) 47-88 (Liss, New York, 1986).
9. Snoek, G. T., Mummery, C. L., Van den Brink, C. E., Van der Saag, P. T. & De Laat, S. W. *Dev. Biol.* **115**, 282-292 (1986).
10. Parker, P. J., Stabel, S. & Waterfield, M. D. *EMBO J.* **3**, 953-959 (1984).
11. Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis (Daudin)* (Elsevier, Amsterdam, 1967).
12. Castagna, M. *et al. J. biol. Chem.* **257**, 7847-7851 (1982).
13. Kraft, A. S. & Anderson, W. B. *Nature* **301**, 621-623 (1983).
14. Davids, M., Loppnow, B., Tiedemann, H. & Tiedemann, H. *Roux's Arch. Dev. Biol.* **196**, 137-140 (1987).
15. Malenka, R. C., Madison, D. V. & Nicoll, R. A. *Nature* **321**, 175-177 (1986).
16. Akers, R. F., Lovings, D. M., Colley, P. A., Linden, D. J. & Routtenberg, A. *Science* **231**, 587-589 (1986).
17. Padgett, R. W., St. Johnston, R. D. & Gelbart, W. M. *Nature* **325**, 81-84 (1987).
18. Hafen, E., Basler, K., Edstroem, J. E. & Rubin, G. M. *Science* **236**, 55-63 (1987).

19. Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. *Cell* **43**, 567-581 (1985).
20. Dickson, C. & Peters, G. *Nature* **326**, 833 (1987).
21. Schackelford, G. M. & Varmus, H. E. *Cell* **50**, 89-95 (1987).
22. Rijsewijk, F. *et al. Cell* **50**, 649-657 (1987).
23. Flickinger, R. A. *J. exp. Zool.* **112**, 165-185 (1949).

Deletion of self-reactive thymocytes occurs at a CD4⁺8⁺ precursor stage

B. J. Fowlkes, R. H. Schwartz & D. M. Pardoll*

Laboratory of Cellular and Molecular Immunology,
National Institute of Allergy and Infectious Diseases,
National Institutes of Health, Bethesda, Maryland 20894, USA

As T cells develop in the thymus, they become tolerant of self-antigens. A major advance in the understanding of how this process occurs was the direct demonstration that cells bearing autoreactive T-cell receptors (TCRs) are physically eliminated from the population of functionally mature T cells present in both the thymus and periphery¹⁻³. We have sought to determine the developmental stage at which autoreactive T cells are eliminated by examining the expression of *V β 17a* anti-I-E TCRs under various experimental conditions. *In vivo* antibody blockage of the CD4 molecule on developing thymocytes of *I-E*⁺ C57BR mice was found to inhibit the deletion of *V β 17a*-bearing cells from the CD4⁺8⁺ single positive thymocyte subset. This result provides strong evidence that deletion of potentially autoreactive T cells occurs at a CD4⁺8⁺ precursor stage, that the non-clonally distributed accessory molecules (CD4, CD8) are significant participants in the self-recognition process that leads to clonal elimination, and that thymic class II major histocompatibility complex (MHC) molecules can influence the repertoire of CD4⁺8⁺ cells.

In F₁ progeny of *V β 17a*⁺ and class II MHC *I-E*⁺ strains, as well as in the C57BR mouse strain (which is simultaneously *I-E*⁺ and *V β 17a*⁺), *V β 17a* is expressed in the functionally incompetent CD4⁺8⁺TCR^{low} thymocyte population⁴⁻⁷, but is almost totally absent in the CD4⁺8⁻ and CD4⁻8⁺ (single positive) thymocyte populations¹. This finding is equally compatible with two opposing theories of T-cell development: (1) that CD4⁺8⁺ thymocytes are a distinct dead-end lineage that is non-functional and, therefore, need not be subject to selection⁸, or (2) that the CD4⁺8⁺ population does contain precursors to mature single positive thymocytes and selection occurs at or after this stage^{9,10}. To experimentally distinguish between these two theories, we focused upon the fact that, in *V β 17a*⁺, *I-E*⁺ mice, *V β 17a*-bearing cells are deleted from the CD4⁺8⁺ as well as the CD4⁺8⁻ subset. One explanation for this finding is that, unlike the majority of anti-class II reactive T cells, TCRs that use *V β 17a* do not require CD4 for anti-I-E reactivity and would, therefore, be autoreactive when expressed by either CD4 or CD8 single positive cells. An alternative explanation is that many *V β 17a*⁺ TCRs do require CD4 for anti-I-E reactivity and that *V β 17a*-bearing cells are deleted from the CD4⁺8⁺ subset because, (as in theory 2), these cells expressed CD4 earlier in their development (that is, they were CD4⁺8⁺) at a time when they were subject to negative selection. A specific prediction of this second explanation is that functionally blocking the CD4 molecule at the CD4⁺8⁺ developmental stage might interfere with I-E recognition and clonal deletion and, thus, result in the appearance of *V β 17a*⁺ TCRs in the CD4⁺8⁺ subset.

To test this prediction, we utilized irradiated C57BR mice reconstituted with syngeneic bone marrow. The kinetics of thymic reconstitution after irradiation and bone marrow transfer have been well documented^{11,12}. Accessory molecules (CD4 and

* To whom reprint requests should be addressed.

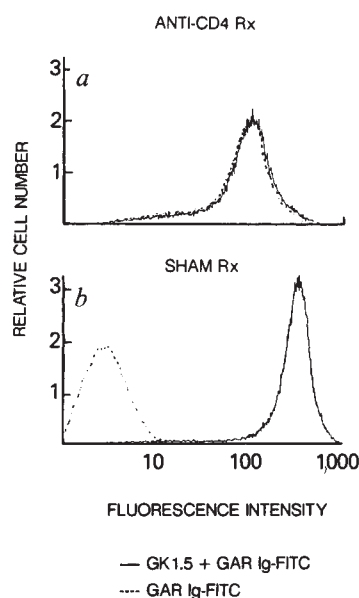


Fig. 1 *In vivo* saturation of CD4 on developing thymocytes in GK1.5-treated C57BR mice. Lethally irradiated C57BR mice were reconstituted with syngeneic C57BR bone marrow. Nine days after bone marrow transfer (BMT), mice were injected intraperitoneally with either GK1.5 MAb¹³ (2 mg every 12 h per mouse, that is 4 mg per day) or equivalent amounts of an isotype-matched control mAb. Thymocytes from the GK1.5-treated (a) and control antibody-treated (b) mice were isolated 23–25 days post-BMT and stained with either GK1.5 + FITC-labelled goat anti-rat IgG second antibody (solid line) or FITC-labelled goat anti-rat IgG second antibody alone (dashed line).

Methods. C57BR mice, obtained from Jackson Laboratories (Bar Harbor), were irradiated with 850R in a Gammacell 40. They were reconstituted with 1×10^7 C57BR bone marrow (BM) cells per mouse and maintained on antibiotic water. GK1.5 MAb ascites was raised in BALB/c mice primed with pristane and treated with Soluocortef (Upjohn) and 500R as described²⁸. GK1.5 was partially purified by precipitation with 40% saturated ammonium sulphate. Effective antibody concentration was measured by titration staining of thymocytes against a purified standard of known concentration. GK1.5 was diluted in saline to a concentration of 10 mg ml⁻¹ in sterile saline, and used at 0.2 ml per injection. The control MAb was a rat anti-rat RT1Bc (Serotech MCA 08) of IgG2b class (same isotype as GK1.5). FITC-labelled goat anti-rat IgG (Caltag) was used to detect *in vivo* bound GK1.5. Fluorescence analysis was performed on a FACS 440 (Becton-Dickinson).

CD8) first appear on donor-derived thymocytes between 10 and 12 days post-transfer. By 23–25 days post-transfer, the thymus consists of >99% donor-derived cells that comprise all four subsets (CD4⁻8⁻, CD4⁺8⁺, CD4⁺8⁻, CD4⁻8⁺) in proportions similar to those of the adult thymus. Thus, by treating syngeneic C57BR chimaeras *in vivo* with the anti-CD4 monoclonal antibody (MAb), GK1.5, (ref. 13) beginning 8 days after bone marrow transfer and analysing reconstituted thymuses at 23–25 days, we could ensure that all thymocytes that expressed CD4 at any point during their development were exposed to the anti-CD4 MAb. Figures 1 and 2 show the effect of an *in vivo* anti-CD4 treatment using the MAb, GK1.5 (2 mg every 12 h per mouse, that is, 4 mg per day) on the major thymocyte subsets. The total number of viable thymocytes recovered from the antibody-treated mice was virtually identical to the sham-injected controls. Staining total thymocytes with GK1.5 followed by fluorescein isothiocyanate-labelled goat anti-rat IgG gave a fluorescence intensity profile identical to that seen on staining with the FITC-labelled goat anti-rat IgG second antibody alone (Fig. 1a), demonstrating complete *in vivo* saturation of surface CD4 molecules with GK1.5. Comparison of CD4 levels on thymocytes from the MAb-treated versus sham-injected mice

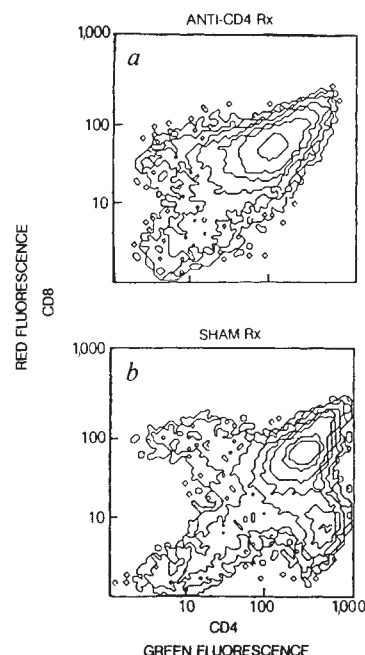


Fig. 2 Effect of *in vivo* anti-CD4 (GK1.5) treatment on thymocyte subsets. Syngeneically BM reconstituted C57BR mice were treated *in vivo* with 4 mg per day GK1.4 (a) or control antibody (b), as in Fig. 1. Thymocytes were removed 23–25 days post-BMT, stained with GK1.5 + FITC-labelled goat anti-rat IgG, followed by CD8.2 (MAb 19/178)²⁹ + Texas red-labelled anti-mouse IgG (Southern Biotech) and by two-colour fluorescence analysis. The slight decrease in CD8 staining seen in panel b is not due to lower expression but, rather, to subsaturating staining with the Texas red labelled IgG.

(Fig. 1a versus 1b) demonstrated a slightly lower intensity of staining which could be the result of *in vivo* modulation of surface CD4. Two-colour analysis of CD4 and CD8 on thymocytes from the GK1.5-treated mice showed that the proportions of CD4⁻8⁻, CD4⁺8⁺, and CD4⁺8⁻ subsets were not significantly affected, whereas the CD4⁻8⁺ subset was almost completely eliminated (Fig. 2a). Reciprocal results (elimination only the CD4⁻8⁺ subset) have been obtained after *in vivo* treatment with anti-CD8 antibodies (data not shown). Therefore, we find that *in vivo* treatment of reconstituting syngeneic chimaeras with antibodies to either accessory molecule (CD4 or CD8): (1) does not detectably eliminate the CD4⁻8⁻ subset; (2) eliminates only the single positive subset expressing the accessory molecule against which the antibody is directed; and (3) binds to, but does not detectably eliminate, CD4⁺8⁺ thymocytes. These results indicate that this experimental design should be suitable for determining whether functional blocking of CD4 on developing CD4⁺8⁺ thymocytes of the C57BR mouse interferes with deletion of *Vβ17a*-expressing cells from the CD4⁺8⁺ subset.

To reliably analyse *Vβ17a* expression, CD4⁻8⁺ thymocytes from the GK1.5-treated syngeneic C57BR chimaeras were enriched by *in vitro* treatment with the MAb J11d (ref. 14) and GK1.5 and complement. This antibody combination not only eliminates many CD4⁻8⁻ cells and virtually all CD4⁺8⁺TCR^{low} cells (which are J11d⁺), it also eliminates a population of non-functional CD4⁻8⁺ thymocytes which express no TCR and are J11d⁺ (refs 15–17). The remaining cells which resist J11d and GK1.5 treatment are CD4⁻8⁺ (all TCR^{bright}) and CD4⁻8⁻. These cells were stained for CD8 and *Vβ17* and analysed by two-colour fluorescence flow cytometry to determine the percentage of CD4⁻8⁺ single positive thymocytes that were *Vβ17a*⁺. The results of one of three such experiments is shown in Fig. 3. While isotype-matched, control antibody-injected C57BR chimaeras expressed virtually no *Vβ17a* in their CD4⁻8⁺ subset (Fig. 3b), the anti-CD4 treated C57BR chimaeras expressed *Vβ17a* on

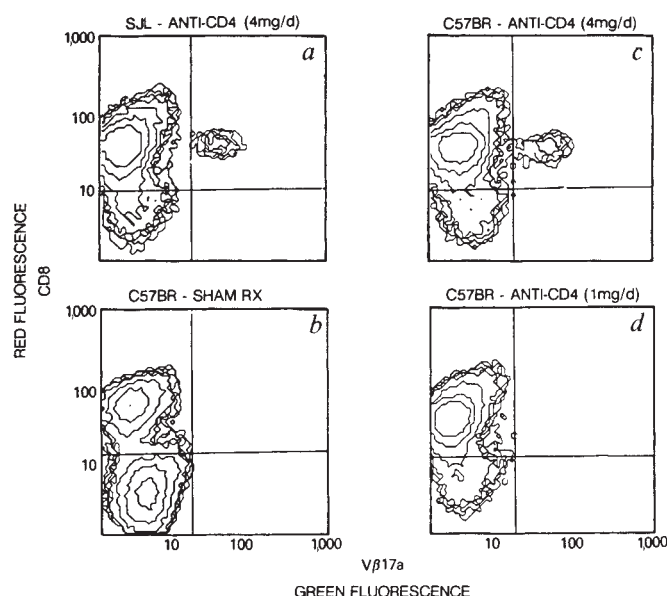


Fig. 3 Appearance of $V\beta 17a^+$ TCR in $CD4^+8^+$ thymocytes after *in vivo* CD4 blockade. Thymocytes from syngeneic bone marrow-reconstituted mice were treated with the MAbs GK1.5+J11d+ complement to enrich for mature cells. The remaining thymocytes were analysed for $V\beta 17a$ expression. The proportions of $CD4^+8^+$ cells expressing $V\beta 17a$ (upper right quadrant/upper right + upper left quadrant) are: 5.8% in 4 mg per day GK1.5-treated SJL ($I-E^-$) mice (a); <0.2% in 4 mg per day control antibody-treated C57BR mice (b); 5.2% in 4 mg per day GK1.5-treated C57BR mice (c); and <0.3% in 1 mg per day GK1.5-treated (resulting in 50% CD4 saturation of thymocytes) C57BR mice (d). The depletion of the left lower quadrant in the anti-CD4 treated mice is due to *in vivo* elimination of $CD4^+8^-$ thymocytes which would fall within this quadrant.

Methods. $V\beta 17a$ staining was performed with KJ23 (ref. 30)+FITC-labelled goat anti-mouse IgG2a (Southern Biotechnology). CD8 staining was performed with a biotinylated anti-Lyt-2 (Becton-Dickinson) followed by allophycocyanin-labelled avidin (Biomeda). Syngeneic BMT, *in vivo* antibody treatments and fluorescence analysis were performed as in Figs 1 and 2.

5.2% of their $CD4^+8^+$ thymocytes (Fig. 3c). This represents a >20-fold increase over the control (Fig. 3b), and is comparable to the frequency of $V\beta 17a$ expression in the $CD4^+8^+$ subset of the $I-E^-$ SJL mouse (Fig. 3a). The finding that $V\beta 17a$ -bearing cells are totally deleted from the $CD4^+8^+$ subset after sub-saturating anti-CD4 treatment (Fig. 3d) indicates that only a portion of CD4 molecules on developing thymocytes are necessary for I-E recognition and subsequent deletion.

The perturbation of the $CD4^+8^+$ TCR repertoire subsequent to *in vivo* anti-CD4 treatment provides strong evidence that $CD4^+8^+$ thymocytes express CD4 at some point in their developmental history, that is their precursors were $CD4^+8^+$. In particular, the appearance of a population of $V\beta 17a^+$ -bearing $CD4^+8^+$ cells in the $I-E^+$ C57BR mouse as a result of anti-CD4 treatment implies that CD4 blockade specifically interfered with clonal deletion of I-E self) reactive thymocytes at a $CD4^+8^+$ precursor stage. This effect of *in vivo* anti-accessory molecule treatment on the single positive subset bearing the alternate accessory molecule is in marked contrast to the results of Smith¹⁸, who found that *in vivo* treatment with an anti-CD8 MAb eliminated the $CD4^+8^-$ subset. It is unlikely that the discrepancy is due to a fundamental difference between anti-CD4 versus anti-CD8 treatment, because we have found that *in vivo* treatment of syngeneic chimaeras with the same anti-CD8 MAb as used in the Smith study resulted in no detectable elimination of either the $CD4^+8^+$ or $CD4^+8^-$ subset (manuscript in preparation). A critical difference in experimental design between these studies is that mixed allogeneic tetraparental

Table 1 Self-tolerance of thymocytes from anti-CD4 treated C57BR mice

Responder*	Stimulator†	c.p.m. per culture	
		Expt 1	Expt 2
C57BR-anti-CD4	C57BR (H-2 ^k)	1,521 ± 51	2,553 ± 101
C57BR-anti-CD4	C57BL/6 (H-2 ^b)	5,074 ± 109	10,065 ± 259
C57BR-anti-CD4	None	1,300 ± 98	2,319 ± 93
C57BR-control	C57BR	2,826 ± 58	4,248 ± 62
C57BR-control	C57BL/6	12,789 ± 946	16,648 ± 644
C57BR-control	None	2,229 ± 37	4,700 ± 232
C57BL/6	C57BR	24,730 ± 2,304	26,662 ± 3,521
C57BL/6	C57BL/6	4,459 ± 227	6,249 ± 994
C57BL/6	None	3,940 ± 363	5,953 ± 988

* Unfractionated thymocytes (2×10^5) were used as responders in experiment 1, and 5×10^4 J11d⁻ thymocytes (enriched for mature, functional thymocytes) were used as responders in experiment 2. The first set of responders were from syngeneic C57BR chimaeras treated with GK1.5 (4 mg per day), as in Fig. 1. The second set of responders were from control syngeneic C57BR chimaeras. The third set of responders (C57BL/6) were used as a control to confirm that irradiated C57BR splenocytes were viable stimulators.

† Irradiated (3,000R) splenocytes (5×10^5) of the indicated strains were used as stimulators for the mixed lymphocyte reactions. MLRs were performed in 96-well round bottom plates. After three days of culture, cells were labelled with ³H-thymidine for 12 h and harvested. IL-2 100 U ml⁻¹ was present during the culture period to enhance any $CD4^+8^+$ responses.

chimaeras were employed in the Smith study, followed by treatment with MAb against the CD8 allele expressed by progeny of one of the transferred bone marrows. It is thus possible that elimination of the $CD4^+8^-$ progeny of the treated parental type was the result of a graft-versus-graft reaction rather than a developmental block.

Further evidence that the CD4 accessory molecule is a crucial co-participant in the anti-I-E reactivity of most developing $V\beta 17a^+$ cells comes from the fact that the $V\beta 17a^+$ $CD4^+8^+$ cells generated in the anti-CD4-treated C57BR mice, although functionally alloreactive to an $I-E^-$ third party (H-2^b) in an *in vitro* mixed lymphocyte response, showed no evidence of anti-self ($I-E^k$) reactivity relative to control treated animals (Table 1). This was seen either when unfractionated or J11d⁻ thymocytes (which are enriched for mature, functional cells) were used as responders. In addition, the demonstration that this subset is not autoreactive argues that the appearance of $V\beta 17a^+$ TCRs among $CD4^+8^+$ thymocytes from anti-CD4 treated C57BR mice is not simply due to elimination of $CD4^+8^-$ cells that are somehow required for maintenance of the thymic tolerance process. Another implication of these results is that negative selection on thymic class II molecules affects the repertoire of the class I-restricted ($CD4^+8^+$) T-cell pool. In the C57BR strain described here, the majority of potential $V\beta 17a$ -bearing, class I-restricted $CD4^+8^+$ thymocytes are eliminated based on cross-reactive I-E (class II) recognition at an earlier $CD4^+8^+$ stage.

The issue of whether $CD4^+8^+$ thymocytes are a dead-end lineage or precursors to functional single positive thymocytes has been difficult to resolve because of the disparate results obtained in attempts to demonstrate differentiation under various conditions. Although some groups have claimed to induce *in vitro* differentiation of thymocyte populations enriched in $CD4^+8^+$ cells¹⁹⁻²², outgrowth of contaminating mature thymocytes, rather than true differentiation, was not ruled out in these studies. In fact, recent studies using $CD4^+8^+$ thymocyte populations highly purified by electronic cell sorting have failed to demonstrate differentiation either after *in vitro* culture²³ or intrathymic transfer (L. Hausmann and B. J. Fowlkes, unpublished). It is likely that the majority of $CD4^+8^+$ thymocytes do, indeed, die within the thymus^{24,25}, either because they fail selection²⁶ or do not synthesize productive TCR molecules²⁷. Thus, if only a small fraction go on to develop into mature single

positives, it might be extremely difficult to directly demonstrate this relatively rare event by analysing whole populations of CD4⁺8⁺ thymocytes.

Despite the difficulty in directly demonstrating differentiation of bulk population CD4⁺8⁺ thymocytes, the experiments presented here argue strongly that thymocytes do pass through a CD4⁺8⁺ stage *en route* to becoming single positives. They further indicate that an important negative selection step occurs during that CD4⁺8⁺ stage in which the accessory molecules participate in the self-recognition process.

We thank Marc Jenkins, Zdenko Kovac and Ron Germain for discussions and Fran Hausman and David Stephany of the NIAID Flow Cytometry Laboratory for fluorescence analysis.

Received 18 April; accepted 15 June 1988.

1. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
2. Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. *Nature* **322**, 35-40 (1988).
3. MacDonald, H. R. *et al. Nature* **332**, 40-45 (1988).
4. Ceredig, R., Glasebrook, A. L. & MacDonald, H. R. *J. exp. Med.* **155**, 358-379 (1982).
5. Chen, W.-F., Scollay, R. & Shortman, K. *J. Immun.* **129**, 18-24 (1982).
6. Harvan, W. L. *et al. Nature* **330**, 170-173 (1987).
7. Finkel, T. H., McDuffie, M. A., Kappler, J. W., Marrack, P. & Cambier, J. C. *Nature* **330**, 179-181 (1987).
8. Shortman, K. & Jackson, H. *Cell. Immun.* **12**, 230-246 (1974).
9. von Boehmer, H. *Immun. Today* **7**, 333-336 (1986).
10. Marrack, P. & Kappler, J. *Science* **238**, 1073-1079 (1987).
11. Sharrow, S. O., Singer, A., Hammerling, U. & Mathieson, B. J. *Transplantation (Baltimore)* **35**, 355-362.
12. Fowlkes, B. J., Edison, L., Mathieson, B. J. & Chused, T. M. *J. exp. Med.* **162**, 802-822 (1985).
13. Dialynas, D. P. *et al. J. Immun.* **131**, 2445-2451 (1983).
14. Bruce, J., Symington, F. W., McKearn, T. J. & Sprent, J. *J. Immun.* **127**, 2496-2501 (1981).
15. Crispe, I. N. & Bevan, M. J. *J. Immun.* **138**, 2013-2018 (1987).
16. Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82-84 (1987).
17. Paterson, D. J. & Williams, A. F. *J. exp. Med.* **166**, 1603-1612 (1987).
18. Smith, L. *Nature* **326**, 798-800 (1987).
19. Irlé, C., Piquet, P.-F. & Vassalli, P. *J. exp. Med.* **148**, 32-45 (1978).
20. Hardt, C., Pfizenmaier, K., Rollinghoff, M., Klein, J. & Wagner, H. *J. exp. Med.* **152**, 1413-1418 (1980).
21. Blue, M.-L., Daley, J. F., Levine, H. & Schlossman, S. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8178-8182 (1985).
22. Kruisbeek, A. M. *Thymus* **1**, 163-185 (1979).
23. Shortman, K. & Scollay, R. in *Recognition and Regulation in Cell-Mediated Immunity* (eds Wiltson, J. D. & Marbrook, J.) 31-60 (Dekker, New York, 1985).
24. Bryant, B. J. *Eur. J. Immun.* **2**, 38-45 (1972).
25. McPhee, D., Pyle, J. & Shortman, K. *Thymus* **1**, 151-162 (1979).
26. Sprent, J., Lo, D., Gao, E.-K. & Ron, Y. *Immunol. Rev.* **101**, 173-190 (1988).
27. Kronenberg, M., Siu, G., Hood, L. E. & Shastri, N. A. *Rev. Immun.* **4**, 529-591 (1986).
28. Weissman, D., Parker, D. J., Rothstein, T. L. & Marshak-Rothstein, A. *J. Immun.* **135**, 1001-1003 (1985).
29. Hammerling, G. J., Hammerling, U. & Flaherty, L. *J. exp. Med.* **150**, 108-116 (1979).

***In vivo* competition between self peptides and foreign antigens in T-cell activation**

Luciano Adorini*, Simone Muller, Francis Cardinaux, Paul V. Lehmann, Fiorenza Falcioni & Zoltan A. Nagy

Preclinical Research, Sandoz Ltd, 4002 Basel, Switzerland

* Laboratory of Pathology, ENEA, CRE Casaccia, 00100 Rome AD, Italy

Cytotoxic and helper T lymphocytes recognize foreign antigen in the form of short peptides associated with class I and class II major histocompatibility complex (MHC) molecules, respectively¹⁻⁴. A recent study of the three-dimensional structure of a class I MHC molecule revealed a cleft formed by the amino-terminal half of the protein, which could serve as the binding site for these peptides⁵⁻⁶. Because an individual possesses only a limited set of different MHC molecules, each molecule of this set must have the ability to bind a large number of different peptides in order to ensure full immunocompetence. Thus, it can be anticipated that peptides with unrelated sequences compete for binding to the same MHC molecule, and, indeed, this has been shown to occur *in vitro*^{7,8}. We therefore decided to see whether such competition could also regulate T-cell responses *in vivo*. We have found that a synthetic peptide corresponding to residues 46-62 of mouse

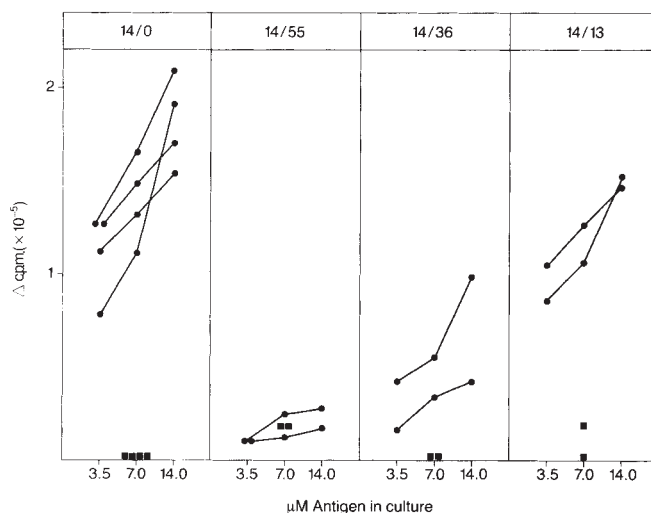


Fig. 1 Peptide ML46-62 inhibits the *in vivo* induction of a T-cell response to peptide HEL46-61 in H-2^k mice. Mice were immunized with HEL46-61 alone or in a mixture with ML46-62 at the antigen/competitor doses (in nmol) indicated at the head of each panel. Lymphocyte proliferative responses to HEL46-61 (●) and ML46-62 (■) were assayed in culture eight days after immunization. The results are expressed as Δ c.p.m., that is, the difference between c.p.m. in cultures with and without antigen and each curve represents the response of an individual mouse. The positive control responses to PPD/ConA, (in Δ c.p.m.) of individual mice (starting from the left panel) were 212,825/452,706; 223,883/390,219; 268,781/250,009; 251,062/155,389; 87,662/424,173; 106,378/241,621; 308,269/288,453; 253,248/343,247; 280,448/351,786; 381,731/289,721. The background values (cells without antigen) ranged from 6,324 to 17,528 c.p.m.

Methods. Peptides were prepared by the solid-phase method on PAM-polystyrene support¹⁸ using side-chain protection, coupling procedures and an automated apparatus (Applied Biosystems, model 430A) as described¹⁹. Peptide resins were subsequently treated with liquid HF at 0°C according to the 'low-high' procedure²⁰, and crude peptides were purified by preparative HPLC on C18 reversed-phase column. The peptides showed correct amino-acid ratios upon hydrolysis in 6N HCl and the expected molecular ions in FAB-MS. Sequences were confirmed by gas-phase microsequencing. Peptide solutions were emulsified in CFA (H37 Ra, Difco, Detroit), and 0.1 ml was injected into the hind footpads of eight-week-old C3H/HeJ mice (SPF, virus-free from Iffa Credo, L'Arbresle, France). After eight days, lymphocytes were prepared from the popliteal lymph nodes and cultured at 4 × 10⁵ per 0.2 ml microtitre well in medium (RPMI 1640, Gibco, Basel) supplemented with 5% human AB serum and different concentrations of HEL46-61 or ML46-62. Responses to purified protein derivative of tuberculin (PPD, Statens Serum-institut, Copenhagen) served as controls for the success of immunization. Proliferation was measured by ³H-thymidine incorporation in triplicate cultures¹¹.

lysozyme, although not immunogenic itself, effectively inhibits the priming for T-cell responses when injected into mice together with foreign protein or peptide antigens. The inhibition observed strictly correlates with the capacity of the competitor to bind to the particular MHC molecule presenting the foreign antigen, and its extent depends on the molar ratio between antigen and competitor.

Peptide 46-61 of hen egg-white lysozyme (HEL) binds to purified I-A^k class II molecules, but does not bind to I-A^d, I-E^d or I-E^k molecules^{2,3}. The immunogenicity of this peptide correlates with its ability to bind to MHC molecules, in that it induces a T-cell response in mice carrying the H-2^k haplotype, but fails to do so in H-2^d mice² (unless 50 fold higher priming doses are used than in H-2^k; Lehmann *et al.*, in preparation). The corresponding murine peptide, mouse lysozyme (ML) 46-62 differs at three amino acid positions from the HEL peptide, and has

Fig. 2 Peptide ML46-62 inhibits the *in vivo* induction of an I-A^k-restricted T-cell response to peptide HEL112-129. The injected antigen/competitor doses are given in nmol in the headings. Response to HEL112-129 of lymphocytes from individual mice (●), and of pooled cells from 2 mice per group (▲) are shown. The positive control responses (PPD/ConA, in Δ c.p.m.) of individual mice (starting from left panel) were: 150,353/69,316; 107,489/55,633; 74,432/71,296; 181,805/67,455; 154,779/49,246; 87,546/117,767; 195,698/162,977; 127,804/187,565; 119,023/116,551; 130,211/128,008. Background values (cells without antigen) ranged from 3,520 to 11,109 c.p.m. Methods as Fig. 1, except B10.A(4R) mice (grade IV isolator reared, Olac) were used.

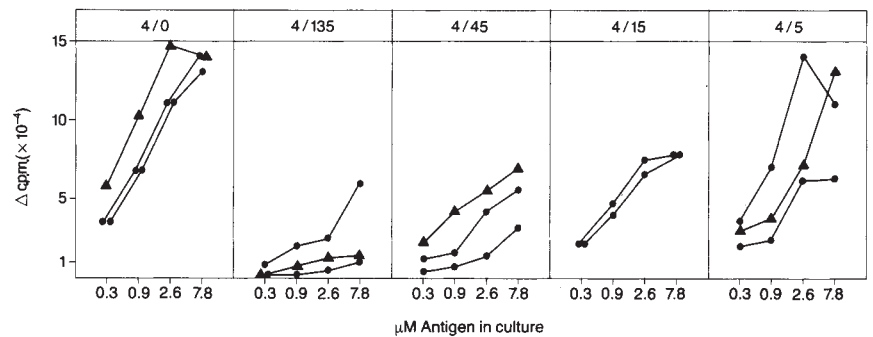


Fig. 3 Influence of peptide ML46-62 on the induction of anti-HEL response in B10.A(4R) mice. The antigen/competitor doses injected are given in nmol in the headings. Responses of individual mice (●) and of pooled cells from 2 mice (▲) to HEL are shown. The positive control responses (PPD/ConA, in Δ c.p.m.) of individual mice (starting from the left panel) were: 126,917/76,313; 159,313/88,322; 95,655/82,688; 139,799/75,250; 166,075/154,276; 112,041/85,904; 81,334/80,303; 192,823/124,039; 105,769/112,119; 184,086/101,449. The background values (cells without antigen) ranged from 3,455 to 13,034.

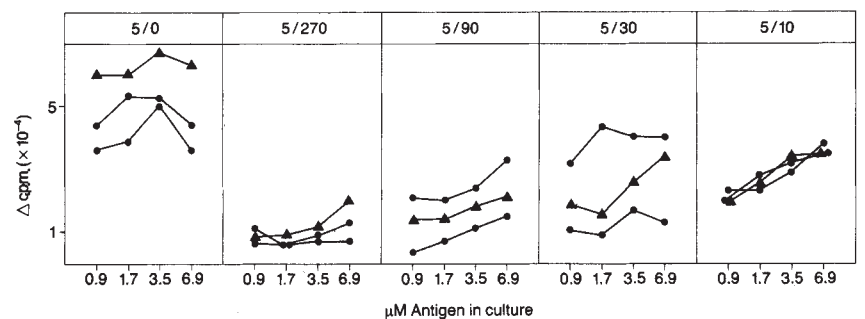
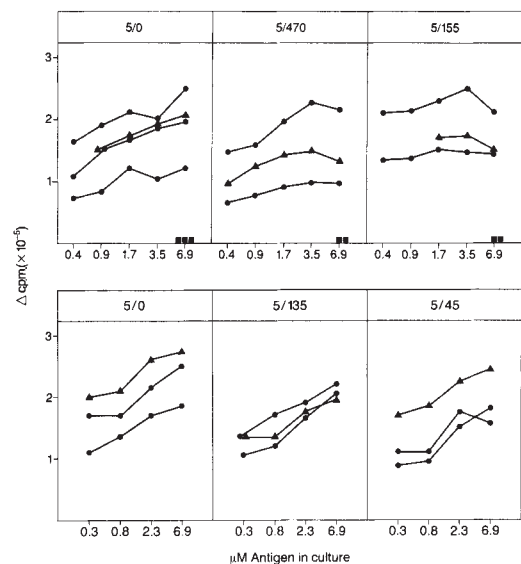


Fig. 4 Failure of peptide ML46-62 to interfere with the induction of T-cell responses to HEL (upper panels) and HEL105-120 (lower panels) in H-2^d mice. The antigen/competitor doses injected are given in nmol in the headings. Response of cells from individual mice to HEL (upper panel, ●), HEL105-120 (lower panel, ●), and ML46-62 (■), and of pooled cells from 2 mice to HEL (upper panel ▲), and HEL105-120 (lower panel ▲) are shown. The positive control responses (PPD/ConA, in Δ c.p.m.) of individual mice (starting from the upper left panel) were: 459,158/366,989; 332,922/278,454; 433,083/315,734; 446,522/278,533; 414,962/312,969; 410,569/226,900; 432,078/178,667; 167,751/198,662; 190,322/249,866; 188,782/351,106; 196,794/195,241; 147,117/294,203; 196,840/338,504. The background values (cells without antigen) ranged from 6,943 to 17,829. Experimental methods as in Fig. 1, except that BALB/c mice (Iffa Credo) were used.



an additional Gly residue at position 48 which is deleted from HEL^{9,10}. The ML46-62 peptide exhibits a similar MHC-binding pattern to the HEL peptide¹⁰, and consequently competes efficiently *in vitro* with immunogenic proteins or peptides for binding to I-A^k molecules of antigen presenting cells (APC), but competes poorly or not at all for I-E^k and I-E^d molecules (ref. 10 and Lehmann *et al.*, in preparation). Furthermore, ML46-62 (as well as ML51-62 which differs only at a single residue from the corresponding HEL peptide) fails to induce a T-cell response in mice even when injected at very high doses (1 mg per mouse), indicating that self-tolerance exists for this peptide (P. V. Lehmann *et al.*, in preparation).

In these experiments we took advantage of the known MHC-

binding characteristics and the lack of immunogenicity of peptide ML46-62, and used it as a competitor during priming for T-cell responses *in vivo*. Groups of mice were immunized with a foreign antigen alone or with mixtures of the antigen and competitor at different molar ratios in complete Freund's adjuvant (CFA). Eight days later, cells from the regional lymph nodes of each individual mouse were tested in a standard T-cell proliferation assay¹¹ for a secondary response to the priming antigen and the competitor. The data in Fig. 1 demonstrate that peptide ML46-62 is capable of inhibiting the induction of a T-cell response to peptide HEL 46-61 in H-2^k mice. A complete inhibition of the response could be observed at an antigen/competitor ratio of ~1/3. The competitor itself was essentially

non-immunogenic. These data establish that two closely related peptides, one immunogenic and the other not, can compete during induction of a T-cell response *in vivo*.

We then tested whether competition can be demonstrated *in vivo* between unrelated peptides binding to the same MHC molecules. We used peptide 112-129 of HEL which, like HEL 46-61, induces an I-A^k-restricted T-cell response (Adorini *et al.*, in preparation), and ML46-62 as a competitor. As shown in Fig. 2, immunization with mixtures of these two peptides results in significantly depressed secondary responses compared to immunization with peptide 112-129 alone. Again the murine peptide was non-immunogenic (data not shown).

It was also of interest to determine whether the ML46-62 peptide could interfere with priming for T-cell responses against the entire HEL molecule. To demonstrate this we used B10.A(4R) mice expressing only the I-A^k class II molecule¹² to which the competitor binds strongly. As shown in Fig. 3, the anti-HEL response of these mice was inhibited by peptide ML46-62 as efficiently as the responses to HEL peptides (Figs 1 and 2).

Finally, we have tested whether the ML46-62 peptide can compete with T-cell responses restricted to MHC molecules to which the competitor binds poorly. Thus we have attempted to inhibit the response of H-2^d mice to HEL, a response consisting of both I-A^d- and I-E^d-restricted components, and also the response to peptide HEL105-120 which induces only a I-E^d-restricted response¹³. The data in Fig. 4 demonstrate a failure of ML46-62 to inhibit either of these responses, indicating that the ability of competitor to bind to class II MHC molecules is a prerequisite for competition to occur *in vivo*.

Collectively, our results demonstrate that the competition between certain peptide antigens for presentation by MHC molecules^{3,4,7,8,14,15} is not only demonstrable *in vitro*, but also occurs on exposure to foreign antigens *in vivo*. The long-standing enigma of antigenic competition¹⁶ may thus, at least in part, be explained by competition for MHC binding sites on APC. The biological significance of competition between self and foreign peptides is not immediately apparent. In principle such a mechanism can be viewed as an obstacle to T-cell activation¹⁷. However, self-nonsel self competition has so far not been observed as a barrier to T-cell activation, indicating that it either does not occur under physiological conditions, or that its effect is to set the lower limit of antigen concentration required for a T-cell response. The latter alternative would make biological sense, because it would reduce 'background' T-cell responses to transiently appearing peptides, such as those derived from pathogens already destroyed by the phagocytic system.

The results also indicate a possibility for immunomanipulation at the level of antigen presentation.

Dissociation of phosphoinositide hydrolysis and Ca²⁺ fluxes from the biological responses of a T-cell hybridoma

Jeffrey J. Sussman*, Mladen Merćep*, Takashi Saito[†], Ronald N. Germain[‡], Ezio Bonvini[‡] & Jonathan D. Ashwell*[§]

* Biological Response Modifiers Program, Division of Cancer Treatment, National Cancer Institute, [†] Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, [‡] Division of Blood and Blood Products, Food and Drug Administration, Bethesda, Maryland 20892, USA

T lymphocytes can be activated in a variety of ways, including occupancy of the T cell antigen receptor (TCR) complex or cross-linking of certain cell-surface molecules with antibody¹⁻⁴. Two of the earliest events seen after stimulation are the hydrolysis of phosphatidylinositol bisphosphate to inositol trisphosphate (Ins P₃) and 1,2-diacylglycerol (DAG)⁵⁻⁷, and an increase in the concentration of intracellular Ca²⁺ ([Ca²⁺]_i)⁸⁻¹¹. Later, the cell secretes lymphokines and expresses lymphokine receptors¹. It has been postulated that the products of the hydrolysis of phosphatidylinositols (Ptd Ins) and fluctuations in [Ca²⁺]_i are critical 'second messengers', transmitting the signals for the initiation of the later events. We have examined the relationship between these second messengers and the secretion of IL-2 in a murine T cell variant whose missing TCR complex had been reconstituted by gene transfer. Surprisingly, although the IL-2 responses of the transfectant could not be distinguished from the original line expressing the same TCR, Ptd Ins hydrolysis and the increase in [Ca²⁺]_i were substantially reduced or absent in the reconstituted cell. It is therefore possible to dissociate these early biochemical changes from a late biological response, raising questions about the putative causal relationship of these events.

We have previously described a spontaneous variant of the 2B4.11 antigen-specific murine T cell hybridoma, designated 21.2.2, that does not express cell surface TCR because it does not transcribe one TCR- α and two TCR- β chain genes¹². Whereas the 21.2.2 cell produced no IL-2 in response to stimulation with antigen, an anti-CD3 antibody, or an anti-Thy-1 antibody, the original 2B4.11 T cell and the TCR- $\alpha\beta$ transfectant, designated Ta β 1.2, responded with virtually identical dose-response curves (Fig. 1a-c). Using larger numbers of the T cell hybridoma, it was possible to detect IL-2 secretion 4-6 hours after stimulation; at these early time points also, 2B4.11 and Ta β 1.2 produced similar (within two-fold) amounts of IL-2 (not shown). To study earlier activation events, the hydrolysis of Ptd Ins to water-soluble inositol phosphates was measured. When co-cultured with concentrations of stimuli that induced maximal IL-2 secretion (see Fig. 1a-c), 2B4.11 cells yielded easily detectable inositol monophosphate (Ins P₁), biphosphate (Ins P₂) and Ins P₃ (Fig. 1), the TCR-negative 21.2.2 cells did not. Surprisingly, although antigen and anti-CD3 did induce detectable Ptd Ins hydrolysis in Ta β 1.2 cells, the inositol phosphate levels achieved were substantially less than in 2B4.11. In comparison with 2B4.11, the maximal amount of Ins P₁ generated by Ta β 1.2 cells in response to anti-CD3 or antigen was 40% and 15%, respectively, while the maximal amount of Ins P₃ generated was only 15% and 7%. Even more striking was the response to the anti-Thy monoclonal antibody (mAb), which induced little or no detectable Ptd Ins turnover in the Ta β 1.2 cells.

Because the failure of Ta β 1.2 cells to exhibit increases in Ptd Ins hydrolysis upon transmembrane stimulation was most clear

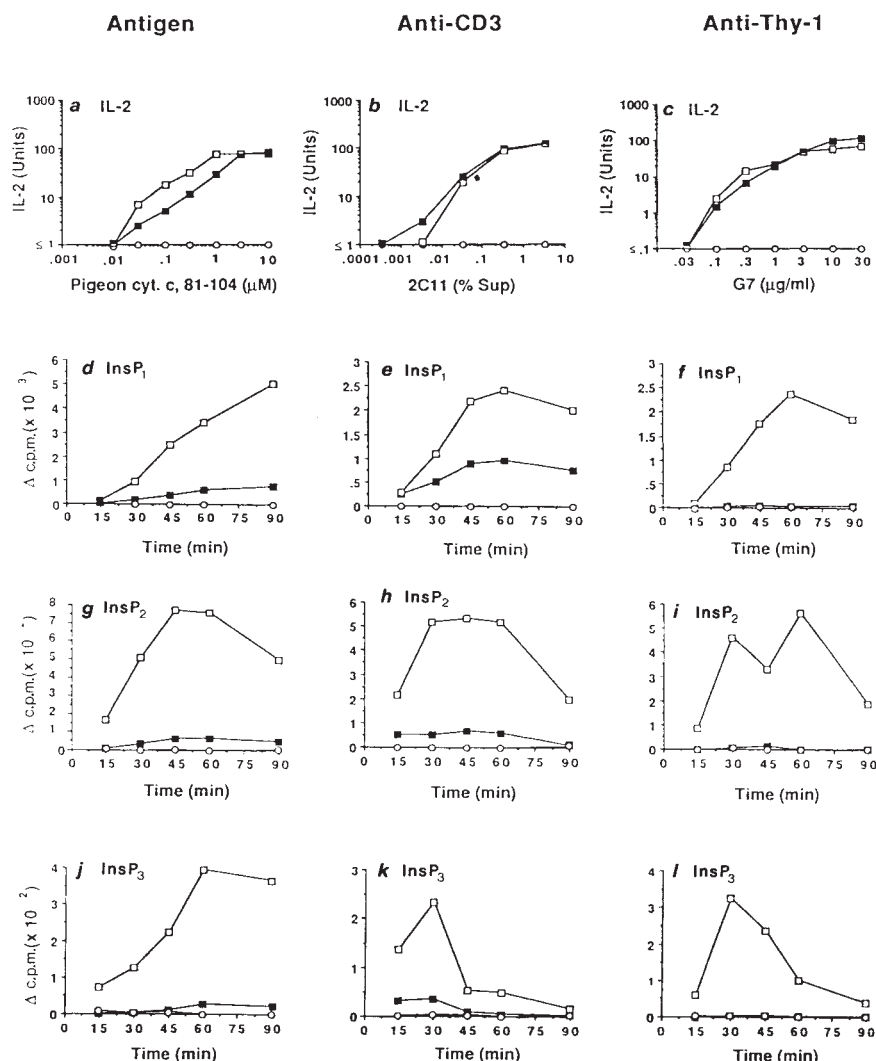
Received 25 February; accepted 14 June 1988.

1. Townsend, A. R. M. *et al.* *Cell* **44**, 959-968 (1986).
2. Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. P. & Unanue, E. R. *Nature* **317**, 359-361 (1985).
3. Buus, S. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 3968-3971 (1986).
4. Guillet, J.-G., Lai, M.-Z., Briner, T. J., Smith, J. A. & Geffer, M. L. *Nature* **324**, 260-262 (1986).
5. Bjorkman, P. J. *et al.* *Nature* **329**, 506-512 (1987).
6. Bjorkman, P. J. *et al.* *Nature* **329**, 512-518 (1987).
7. Guillet, J.-G. *et al.* *Science* **235**, 865-870 (1987).
8. Buus, S., Sette, A., Colon, S. M., Miles, C. P. & Grey, H. M. *Science* **235**, 1353-1358 (1987).
9. Riblet, R. L. in *Lysozyme* (eds Osseman, E. F., Canfield, R. E. & Beychoke, S.) 89-93 (Academic, New York, London, 1974).
10. Babbitt, B. P., Matsueda, G., Haber, E., Unanue, E. R. & Allen, P. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4509-4513 (1986).
11. Corradin, G., Etlinger, H. M. & Chiller, J. M. *J. Immun.* **119**, 1048-1053 (1977).
12. Jones, P. P., Murphy, D. B. & McDevitt, H. O. *J. exp. Med.* **148**, 925-939 (1978).
13. Adorini, L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
14. Werdelin, O. *J. Immun.* **129**, 1883-1891 (1982).
15. Rock, K. L. & Benacerraf, B. *J. exp. Med.* **157**, 1618-1634 (1983).
16. Radovich, J. & Talmage, D. W. *Science* **158**, 512-514 (1967).
17. Raff, M. C. *Nature* **298**, 791-792 (1982).
18. Mitchell, A. R., Erickson, B. W., Ryabitshev, M. N., Hodges, R. S. & Merrifield, R. B. *J. Am. chem. Soc.* **98**, 7357-7362 (1976).
19. Clark-Lewis, I., Aebersold, R., Ziltener, H., Schrader, J. W., Hood, L. E. & Kent, S. B. H. *Science* **231**, 134-139 (1986).
20. Tam, J. P., Heath, W. F. & Merrifield, R. B. *J. Am. chem. Soc.* **105**, 6442-6455 (1983).

§ To whom correspondence should be addressed.

Fig. 1 Stimulation of IL-2 and water-soluble inositol phosphate production by antigen, anti-CD3 mAb, or anti-Thy-1. Aliquots of 2B4.11 ($\square$), 21.2.2 ($\circ$), or $\text{T}\alpha\beta$ 1.2 cells ($\blacksquare$) were stimulated with antigen (pigeon cytochrome *c* fragment 81-104; ref. 34), anti-CD3 (2C11; ref. 33, kindly provided by Dr Jeffrey Bluestone, Ben May Institute, University of Chicago), or anti-Thy-1 (G7; ref. 2). *a-c*, IL-2 production; *d-l*, generation of phosphoinositides.

Methods. IL-2 production: 10^4 T-cell hybridomas per well were stimulated with antigen, 2C11, or G7 for 20 h, and an aliquot of supernatant was removed for measurement of IL-2 as described¹². One unit of IL-2 is defined as that dilution of supernatant that stimulates an IL-2-dependent T cell to half of its maximal response. Measurement of Ptd Ins hydrolysis was modified from the procedure of Berridge³⁶. Briefly, 2×10^7 cells ml^{-1} were loaded with $20 \mu\text{Ci ml}^{-1}$ of myo-[^3H (N)]-inositol (^3H -myo-inositol, 14 Ci mmol^{-1} ; NEN) in inositol-free medium for 3 h at 37°C , and washed with medium containing 10 mM LiCl. The three cell types were comparable in their ability to take up [^3H]-myo-inositol and incorporate it into lipid-soluble compartments. To determine the specific activity of the labelled Ptd Ins pool, lipids were extracted and analysed by TLC. The Ptd Ins band was scraped and the phospholipids hydrolysed with sulphuric acid (155°C for 18 h) and analysed for phosphorus and radioactivity content. The specific activity of ^3H was 34.8 pCi pmol^{-1} of Ptd Ins for 2B4.11 cells, and 30.8 pCi pmol^{-1} of Ptd Ins for $\text{T}\alpha\beta$ 1.2 cells. To assess the degree of Ptd Ins hydrolysis, 10^6 labelled T cells per experimental point were added in duplicate to polypropylene tubes containing either 10^6 LK 35.2 cells³⁵ (as antigen-presenting cells) with or without fragment 81-104 ($10 \mu\text{M}$ final), 10^6 LS 102.9 cells³⁵ (as a source of Fc receptors to cross-link the mAb³²) with or without 2C11 supernatant (3.3% final), or medium with or without G7 ($30 \mu\text{g ml}^{-1}$ final). Stimulation at 37°C was stopped at the times indicated by adding a mixture of chloroform:methanol, and the aqueous phase applied to an anion exchange resin column (AG 1-X8, formate form; Bio-Rad Laboratories). The different inositol phosphate fractions were sequentially eluted with NH_4 formate/formic acid buffers of increasing concentration³⁶. The 1M NH_4 formate/0.1M formic acid eluate, designated Ins P_3 , probably includes small amounts of Ins P_4 as well. To calculate the change in c.p.m., the c.p.m. achieved at each time point in the presence of medium alone was subtracted from that obtained with the stimuli.



with the anti-Thy-1 mAb, subsequent studies were performed with G7. A different technique, measurement of incorporation of ^{32}P into phospholipids over time, was used to confirm independently that Ptd Ins hydrolysis in response to transmembrane stimuli was deficient in $\text{T}\alpha\beta$ 1.2. Both 2B4.11 and $\text{T}\alpha\beta$ 1.2 cells were loaded with ^{32}P and incubated in the presence or absence of the anti-Thy-1 mAb (Fig. 2). The amount of label incorporated into the different lipid fractions was determined at the intervals shown by the thin layer chromatography (TLC). After exposure to anti-Thy-1 the amount of ^{32}P label incorporated into Ptd Ins in 2B4.11 increased steadily, reflecting the movement of labelled precursors into newly-synthesized Ptd Ins (Fig. 2a). No such increase was observed in the $\text{T}\alpha\beta$ 1.2 cells (Fig. 2d). Phosphatidic acid is synthesized by ATP-dependent phosphorylation of DAG, and its levels reflect the amount of DAG substrate present. DAG is generated by the phospholipase C-mediated hydrolysis of phospholipids, including phosphoinositides, and is a physiological activator of protein kinase C (PKC). In 2B4.11 cells the labelling of phosphatidic acid was found to increase rapidly and progressively in response to anti-Thy-1, whereas in $\text{T}\alpha\beta$ 1.2 cells there was relatively little incorporation of ^{32}P into phosphatidic acid and no progression with time (Fig. 2b and e). In both cells the labelling of phosphatidylethanolamine and phosphatidylcholine was unaffected or only

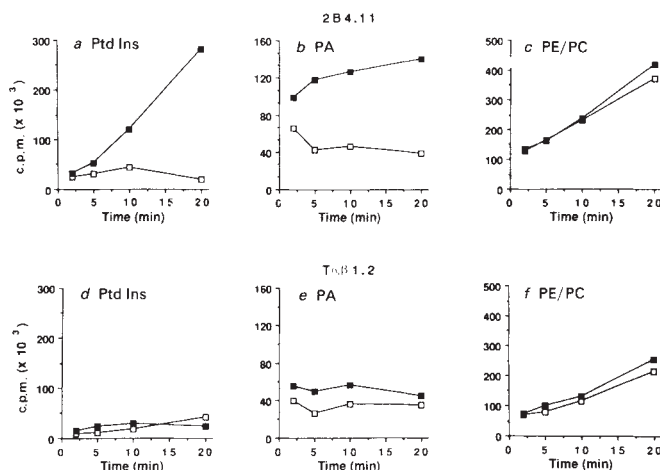
slightly increased after treatment with G7 (Fig. 2c and f). The small elevation in phosphatidic acid that was noted in $\text{T}\alpha\beta$ 1.2 cells may have been due to undetectable levels of Ptd Ins hydrolysis. Alternatively, it may reflect a small amount of DAG generated in Ptd Ins independent pathways, as has been reported for a variety of cell types after stimulation with platelet-derived growth factor or phorbol esters¹³, and for mast cells after cross-linking of surface IgE receptors¹⁴. In any case, anti-Thy-1-induced incorporation of ^{32}P into phosphatidic acid was substantially diminished in $\text{T}\alpha\beta$ 1.2 cells, supporting the finding that $\text{T}\alpha\beta$ 1.2 cells generated little inositol phospholipid hydrolysis after stimulation with G7.

Two metabolites of Ptd Ins hydrolysis, inositol 1,4,5-trisphosphate and inositol 1,3,4,5-tetrakisphosphate, have been implicated in increasing $[\text{Ca}^{2+}]_i$ either by releasing Ca^{2+} from intracellular stores or by opening plasma membrane Ca^{2+} channels¹⁵⁻¹⁹. Therefore, $[\text{Ca}^{2+}]_i$ was measured in the three cell types before and after stimulation with anti-Thy-1. Activation of 2B4.11 with G7 resulted in an $\sim 300\%$ increase in the level of $[\text{Ca}^{2+}]_i$ (Fig. 3). Consistent with their failure to generate Ins P_3 , however, neither the 21.2.2 nor $\text{T}\alpha\beta$ 1.2 cells exhibited a change in $[\text{Ca}^{2+}]_i$ upon activation.

The quantitatively normal production of IL-2 in the absence of easily detectable PTD Ins turnover or changes in $[\text{Ca}^{2+}]_i$,

Fig. 2 Effect of an anti-Thy-1 mAb on the incorporation of [32 P]-orthophosphate into phospholipids. [32 P]-orthophosphate-labelled 2B4.11 cells (a–c) or $\text{T}\alpha\beta 1.2$ cells (d–f) were exposed to medium ($\square$) or 1:50 final dilution of G7 ascites ($\blacksquare$). At the indicated times, the incubation was terminated and lipids extracted. Incorporation of ^{32}P into a, c Ptd Ins; b, d, phosphatidic acid; and c, f, the combination of phosphatidylethanolamine and phosphatidylcholine (PE/PC), as determined by TLC, is shown.

Methods. Aliquots of $25\text{--}40 \times 10^6$ cells were washed twice with a saline solution containing 25 mM HEPES, pH 7.35 (M. A. Bio-products, Walkersville, MD) and 1 mg ml $^{-1}$ of bovine serum albumin (BSA, Sigma), and depleted of intracellular phosphate by incubation for 30 min at 37°C in 10 ml of RPMI 1640 (Gibco) containing 25 mM HEPES, 10% fetal calf serum that had been dialysed against saline, and 0.1 mM sodium phosphate ('low phosphate medium'). At the end of the incubation, cells were washed and resuspended in 1 ml of low phosphate medium supplemented with 0.75 to 1 mCi of [32 P]-orthophosphate (NEN). After labelling for 1 h at 37°C, the cells were washed three times and resuspended at a concentration of $6\text{--}8 \times 10^6$ cells ml $^{-1}$ in low phosphate medium lacking phenol red. The cells ($1.5\text{--}2 \times 10^6$ cells in 250 μl aliquots) were dispensed into 15 ml polypropylene tubes, and 250 μl of medium or G7-containing ascites was added. The cell suspensions were incubated at 37°C, and incubations were stopped by the addition of 1.5 ml of a mixture of water-saturated chloroform:methanol 1:2 (v:v) followed by immediate vortexing. Lipids were extracted by phase partition with the addition of 0.5 ml of chloroform and 0.5 ml of 0.1 N hydrochloric acid, using a modification of the method described by Folch *et al.*³⁷. The chloroform phase was recovered, washed once with an equal volume of methanol:0.1 N hydrochloric acid:phosphate buffered saline (2:1:1, v:v:v), passed through 0.2 μm glass filters (Schleicher and Schuell), and an aliquot counted for the amount of radioactivity incorporated. The remainder was evaporated to dryness under nitrogen and re-dissolved in a small amount of chloroform:methanol, 2:1 (v:v). Samples were placed onto pre-lined silica-gel thin-layer chromatography plates (LK5 20 $\times$ 20 cm plates, Whatman), pre-impregnated with 2% (w:v) boric acid solution. To facilitate comparison between cells, an equal number of c.p.m. was applied to each lane. Plates were first developed for half their length with chloroform in saturated tanks, and after drying, developed for their full length using chloroform:methanol:4 M ammonium hydroxide, 75:58:17 (v:v:v), as the solvent system³⁸. Authentic standards were visualized by iodine vapor staining. The labelled phospholipids were detected by autoradiography. The relative amount of label incorporated in each phospholipid species was measured by densitometry scanning of the autoradiograms using an LKB UltroScan XL laser densitometer. The absolute (total) amount of radioactivity incorporated per phospholipid species was computed from the relative amount and the total radioactivity recovered.



might reflect an unusually high degree of sensitivity to these biochemical messengers in $\text{T}\alpha\beta 1.2$ cells. To examine this, IL-2 production was measured after stimulation with chemicals thought to mimic DAG and InsP_3 , 12-*O*-tetradecanoyl phenol 13-acetate (TPA) (an activator of PKC) and ionomycin (a Ca^{2+} ionophore) (Fig. 4). The two cell lines were indistinguishable in their secretion of IL-2. To test the possibility that in $\text{T}\alpha\beta 1.2$ cells undetectably small changes in $[\text{Ca}^{2+}]_i$ were sufficient to result in IL-2 secretion, $[\text{Ca}^{2+}]_i$ was measured in $\text{T}\alpha\beta 1.2$ in the presence of fetal calf serum, 10 ng ml $^{-1}$ of TPA, and ionomycin. At 0.1 μg ml $^{-1}$ of ionomycin (the lowest concentration at which IL-2 was detected) a rapid rise in $[\text{Ca}^{2+}]_i$ was easily observed, increasing by ~ 10 -fold over the baseline level (data not shown). Together these experiments suggest that normal IL-2 production by $\text{T}\alpha\beta 1.2$ cells in the absence of detectable changes in Ptd Ins metabolism or $[\text{Ca}^{2+}]_i$ is not due to an unusually low threshold for a response to these second messengers.

In lymphocytes, the potential mediators of intracellular signal transmission that have received particular attention are DAG, inositol phosphates (especially InsP_3 and InsP_4), and fluctuations in the level of $[\text{Ca}^{2+}]_i$, each of which has been shown to increase rapidly in response to transmembrane stimuli^{7,10,11,17}. The current paradigm that Ptd Ins hydrolysis and changes in $[\text{Ca}^{2+}]_i$ are critical events in T-cell activation comes from the temporal correlation between changes in their levels and the subsequent secretion of IL-2, and the observation that simultaneously activating PKC and increasing $[\text{Ca}^{2+}]_i$ with chemical reagents generally leads to the late consequences of activation^{20,21}. Furthermore, removal of extracellular Ca^{2+} inhibits activation-induced IL-2 secretion²²⁻²⁴. Therefore, in studying the effect of TCR expression loss on T cell activation, we anticipated that changes in the capacity of a cell to modify its resting state of Ptd Ins hydrolysis and $[\text{Ca}^{2+}]_i$ in response to membrane perturbations would vary in parallel with its capacity to produce late biological responses. The clear but unexpected finding that this was not the case for $\text{T}\alpha\beta 1.2$ cells may be accounted for in one of two ways. First, the products of Ptd Ins

hydrolysis and increases in $[\text{Ca}^{2+}]_i$ may not be the relevant second messengers in $\text{T}\alpha\beta 1.2$ (and by extension 2B4.11) that are responsible for IL-2 production; fluctuations in their levels might or might not be involved in the initiation of biological responses that we have not measured. Second, it may be that increases in InsP_3 and $[\text{Ca}^{2+}]_i$ are necessary to signal for IL-2 production, but that the levels required (and generated in these cells) are below the limits of our ability to detect them. The finding that concentrations of anti-Thy-1 mAb that yielded maximal IL-2 production and inhibition of transformed growth¹² failed to elicit detectable changes in either Ptd Ins hydrolysis or $[\text{Ca}^{2+}]_i$ in $\text{T}\alpha\beta 1.2$ cells implies that the levels generally measured under experimental conditions have little relevance for these late sequelae of activation.

The reason for the failure of $\text{T}\alpha\beta 1.2$ cells to generate detectable amounts of inositol phosphates in response to activating stimuli is unknown. We have observed that in comparison to 2B4.11, $\text{T}\alpha\beta 1.2$ produces equivalent amounts of InsP_1 and approximately half as much InsP_2 and InsP_3 in response to AlF_4^- , which is thought to stimulate directly by a guanine nucleotide-binding (G) protein²⁵⁻²⁸. We conclude that although an AlF_4^- -sensitive system at the level of, or distal to, G protein involvement might be partially defective in $\text{T}\alpha\beta 1.2$, such a defect is unlikely to be responsible for the complete lack of detectable Ptd Ins hydrolysis in response to anti-Thy-1. Another possibility is that $\text{T}\alpha\beta 1.2$ cells are deficient in phospholipase C, or that this enzyme fails to interact normally with the TCR complex and/or its associated signalling structures. The ability of AlF_4^- to induce some Ptd Ins hydrolysis in $\text{T}\alpha\beta 1.2$ cells suggests that a G protein-inducible phospholipase C is active, but whether it is the same enzyme that is responsible for Ptd Ins hydrolysis upon stimulation via surface membrane molecules is not known.

The lack of correlation observed in $\text{T}\alpha\beta 1.2$ cells between early and late activation events usually considered to be cause and effect may not be rare. We have found that EL-4, although it produces large amounts of IL-2 in response to G7, fails to manifest increases in $[\text{Ca}^{2+}]_i$ (not shown). Furthermore, after

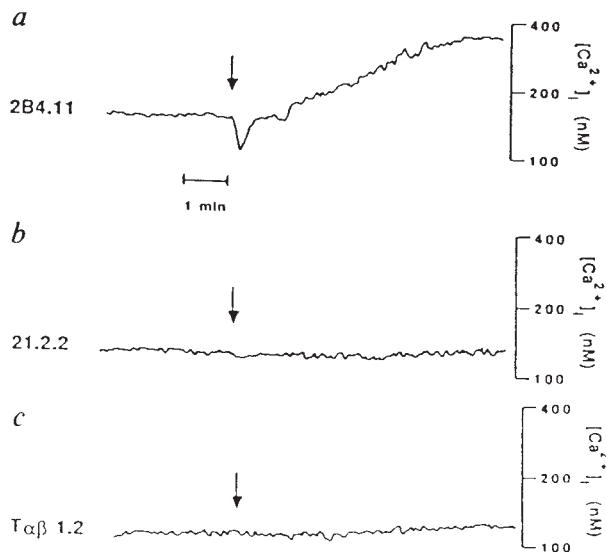


Fig. 3 Changes in $[Ca^{2+}]_i$ in response to an anti-Thy-1 mAb. Cells from the 2B4.11, 21.2.2, or $T\alpha\beta 1.2$ lines were loaded with the fluorescent indicator fura-2 and analysed for fluctuations in $[Ca^{2+}]_i$ after the addition of G7.

Methods. T cells (10^6 per ml) were incubated at $37^\circ C$ for 30 min in the presence of $2 \mu M$ of fura-2/AM (Calbiochem). The dye-loaded cells were washed in phosphate-buffered saline and resuspended in a solution of 140 mM NaCl, 5 mM KCl, 1 mM $MgSO_4$, 1 mM $CaCl_2$, 20 mM HEPES, 1 mM NaH_2PO_4 , 5.5 mM glucose, pH 7.4, supplemented with 5% fetal calf serum. Aliquots of 4×10^6 cells in 2 ml were placed into a fluorimeter (LS-5, Perkin-Elmer, Norwalk, CT) in magnetically stirred cuvettes in which the temperature was maintained at $35^\circ C$ (to reduce leakage of fura-2)³⁹. The cells were excited at a wavelength of 340 nm and emission monitored at 505 nm. Once a baseline was established, G7 ascites (1:500 final concentration) was added and the cells monitored for 7 min. The fura-2 fluorescence signal was calibrated for $[Ca^{2+}]_i$ as described³⁹. Briefly, at the end of each experiment, the contribution of extracellular fura-2 to the total fluorescence signal was determined by adding 2 mM EGTA and 40 mM Tris, and monitoring the initial rapid fall of the fluorescence. F_{min} was determined by adding 0.025% Triton X-100 in the presence of the EGTA, and F_{max} was obtained by adding back 4 mM $CaCl_2$ to the same sample. After correcting for the contribution made by extracellular Ca^{2+} , $[Ca^{2+}]_i$ was calculated according to the formula: $[Ca^{2+}]_i = K_d[(F - F_{min})/(F_{max} - F)]$ where K_d is $225 nM$ ⁴⁰.

preincubation with a phorbol ester and the mitogen phytohaemagglutinin (PHA) in the absence of extracellular Ca^{2+} , whole T-cell populations will proceed to proliferate²⁹, and a lack of correlation has been observed between anti-immunoglobulin stimulated B-cell proliferation and changes in Ptd Ins hydrolysis and $[Ca^{2+}]_i$ ³⁰. Beaven and colleagues have described a variant of the RBL-2H3 basophilic leukaemia cell that secretes histamine and hydrolyses Ptd Ins but fails to manifest an increase in $[Ca^{2+}]_i$ in response to antigen³¹. Considered together with the data in the present report, these findings dissociate experimentally detectable generation of certain biochemical second messengers from the initiation of particular activation-dependent cellular responses. Our results raise the possibility that the measurable changes in levels of Ins P_3 and $[Ca^{2+}]_i$ that are usually observed soon after stimulation, and which were clearly evident in the original 2B4.11 cells, are not quantitatively, or perhaps even qualitatively, essential events in the pathway(s) leading to IL-2 secretion. If so, the signal transduction mechanism in these T cells that is relevant for the initiation of this late activation event remains to be determined.

We thank Drs Michael Beaven and Ronald Schwartz for critical review of this manuscript. J.J.S. is the recipient of a Howard Hughes Research Scholar Award.

Note added in proof: Recent experiments have indicated that

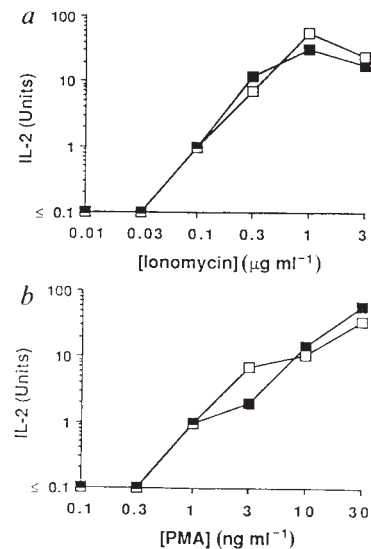


Fig. 4 Stimulation with TPA and ionomycin. Aliquots of 10^4 2B4.11 (□) or $T\alpha\beta 1.2$ cells (■) were incubated with a, $10 ng ml^{-1}$ of TPA and varying concentrations of ionomycin or b, $1 \mu g ml^{-1}$ of ionomycin and varying doses of TPA. After 24 h, supernatant was removed and assayed for IL-2. In b, $1 \mu g ml^{-1}$ of ionomycin alone induced 2 units of IL-2 from 2B4.11 and 3 units from $T\alpha\beta 1.2$; these values were subtracted before plotting the data

$T\alpha\beta 1.2$ cells are deficient in the CD3- η chain, a phenotype that correlates with decreased Ptd Ins hydrolysis in a number of independently-derived variants of the 2B4.11 T cell hybridoma (M. Mercep *et al.*, manuscript in preparation).

Received January 7; accepted 14 June 1988.

- Ashman, R. F. in *Fundamental Immunology* (Raven, New York, 1984).
- Gunter, K. C., Malek, T. R. & Shevach, E. M. *J. exp. Med.* **159**, 716-730 (1984).
- Malek, T. R., Ortega, G., Chan, C., Kroczyk, R. A. & Shevach, E. M. *J. exp. Med.* **164**, 709-722 (1986).
- Meuer, S. C. *et al. Cell* **36**, 897-906 (1984).
- Resch, K., Gelfand, E. W., Hansen, K. & Ferber, E. *Eur. J. Immun.* **2**, 598-601 (1972).
- Hui, D. Y. & Harmony, J. A. K. *Biochem. J.* **192**, 91-98 (1980).
- Imboden, J. B. & Stobo, J. D. *J. exp. Med.* **161**, 446-456 (1985).
- Allwood, G., Asherson, G. L., Davey, M. J. & Goodford, P. J. *Immun.* **21**, 509-516 (1971).
- Whitney, R. B. & Sutherland, R. M. *Cell Immun.* **5**, 137-147 (1972).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68-71 (1982).
- Weiss, A., Imboden, J., Shoback, D. & Stobo, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4169-4173 (1984).
- Sussman, J. J., Saito, T., Shevach, E. M., Germain, R. N. & Ashwell, J. D. *J. Immun.* **140**, 2520-2526 (1988).
- Besterman, J. M., Duronio, V. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6785-6789 (1986).
- Kennerly, D. A. *J. biol. Chem.* **34**, 16305-16313 (1987).
- Berridge, M. J. *Biochem. J.* **220**, 345-360 (1984).
- Whitaker, M. & Irvine, R. R. *Nature* **312**, 636-639 (1984).
- Stewart, S. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 6098-6102 (1986).
- Irvine, R. F. & Moor, R. M. *Biochem. J.* **240**, 917-920 (1986).
- Kuno, M. & Gardner, P. *Nature* **326**, 301-304 (1987).
- Weiss, A. & Stobo, J. D. *J. exp. Med.* **160**, 1284-1299 (1984).
- Truneh, A., Albert, F., Golstein, P. & Schmitt-Verhulst, A.-M. *Nature* **313**, 318-320 (1985).
- Alford, R. H. *J. Immun.* **104**, 698-703 (1970).
- Whitney, R. B. & Sutherland, R. M. *J. cell Physiol.* **80**, 329-338 (1972).
- Freedman, M. H. *Cell Immun.* **44**, 290-313 (1979).
- Sternweis, P. C. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4888-4891 (1982).
- Blackmore, P. F., Boeckino, S. B., Waynick, L. E. & Exton, J. H. *J. biol. Chem.* **260**, 14477-14483 (1985).
- Strnad, C. F. & Wong, K. *Biochem. Biophys. Res. Commun.* **133**, 161-167 (1985).
- O'Shea, J. J. *et al. J. Immun.* **139**, 3463-3469 (1987).
- Gelfand, E. W., Cheung, R. K., Mills, G. B. & Grinstein, S. *Nature* **315**, 419-420 (1985).
- Roifman, C. M. *et al. Eur. J. Immun.* **17**, 1737-1742 (1987).
- Woldemussie, E., Ali, H., Takaishi, T., Siraganian, R. P. & Beaven, M. A. *J. Immun.* **139**, 2431-2438 (1987).
- Mercep, M., Bluestone, J. A., Noguchi, P. D. & Ashwell, J. D. *J. Immun.* **140**, 324-335 (1988).
- Leo, O., Foo, M., Sachs, D. H., Samelson, L. E. & Bluestone, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1374-1378 (1987).
- Corradin, G. & Harbury, H. A. *Biochim. biophys. Acta* **221**, 489-496 (1970).
- Kappler, J., White, J., Wegmann, D., Mustain, E. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3604-3607 (1982).
- Berridge, M. J. *Biochem. J.* **212**, 849-858 (1983).
- Folch, J., Lees, M. & Sloane-Stanley, G. H. *J. biol. Chem.* **226**, 497-509 (1957).
- Fine, B. J. & Sprecher, H. *J. Lipid Res.* **23**, 660-663 (1982).
- Treves, S. *et al. J. exp. Med.* **166**, 33-42 (1987).
- Gryniewicz, G., Ponnie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440-3450 (1985).

Induction of proto-oncogene *JUN/AP-1* by serum and TPA

William W. Lamph, Penny Wamsley,
Paolo Sassone-Corsi & Inder M. Verma

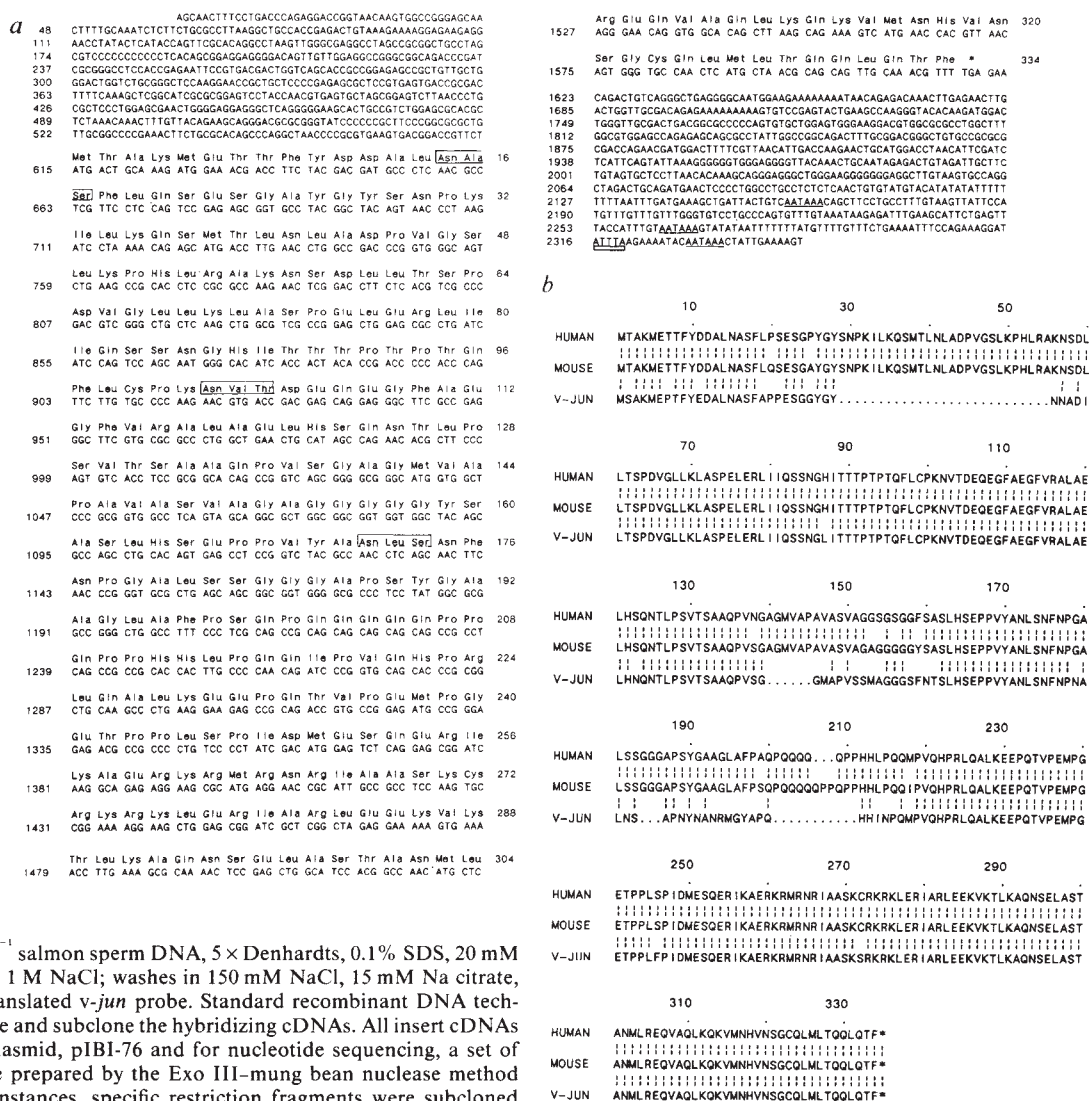
Molecular Biology and Virology Laboratory, The Salk Institute,
PO Box 85800, San Diego, California 92138, USA

The response of a cell to mitogens and differentiation agents involves the transcriptional induction of several cellular genes. Prominent among these so-called 'immediate early' or 'competence' genes are the nuclear oncogenes *fos* and *myc*¹⁻⁷. Although the precise function of these early response genes in growth control is not understood, it is likely that many of them are involved in the transition from G₀ to G₁ in the cell cycle. The findings that the products of nuclear proto-oncogenes *jun* and *erbA* are transcription factors supports the notion of the role of the nuclear oncoproteins in the regulation of gene expression⁸⁻¹¹. Recently, it has been reported that the FOS protein is associated in transcriptional complexes with the product of the *jun* oncogene, the transcription factor AP-1 (refs 12-15). As the *fos* gene is induced in response to mitogens during initiation of cell growth, we investi-

gated whether expression of the nuclear transcription factor AP-1 is also inducible. We report that mouse *c-jun* gene transcription is rapidly induced by serum and phorbol-ester 12-*o*-tetradecanoyl phorbol 13-acetate (TPA). Furthermore, induction is transient and the mRNA is superinduced by inhibitors of protein synthesis.

AP-1 is a transcription factor first identified as a protein that recognized a specific sequence in the *cis*-control regions on the SV40 virus and the human metallothionein IIA gene^{16,17}. The AP-1 binding site, TGACTCA, has subsequently been identified in a large number of other genes, most of which have been shown to be inducible with tumour promoter phorbol-esters. The human AP-1 gene has been shown to be the cellular homologue of the transforming *v-jun* gene of an avian sarcoma virus (ASV-17)^{8,9}. We isolated several complementary DNA clones corresponding to the murine cellular homologue of *v-jun*. The longest cDNA clone obtained (pcD10) is 2,347 nucleotides long which is in close agreement with the size of the mouse *c-jun* messenger RNA (2.7 kilobase (kb), see Fig. 2a). Three additional clones are co-linear with this sequence, but are truncated at their 5'-ends, probably due to incomplete extension during DNA synthesis. The nucleotide sequence is shown together with the deduced amino-acid sequence in Fig. 1a. A long (614 base pair) 5'-untranslated region contains termination codons in all three frames. The first methionine start codon is at position 615 and corresponds to the first methionine in the

Fig. 1 Nucleotide and predicted amino-acid sequence of murine *c-jun* cDNA. **a**, The sequence of the 5'-untranslated region, the coding domain and the 3' non-coding region are shown. Nucleotide positions are shown on the left and the deduced amino-acid positions on the right. Important landmarks include three poly(A) addition signal sequences (—), the ATTTA putative destability sequence (□) and possible N-linked glycosylation sites (□). **b**, Amino-acid-sequence identities between murine human and viral JUN proteins. The open reading frame of murine *c-jun* cDNA is compared with the previously published amino-acid sequences of human c-JUN and v-JUN proteins^{9,19}. The program of Smith and Waterman was used²⁴. Identities are marked. Gaps indicate absence of the amino acid. **Methods.** A eukaryotic cDNA expression library, prepared from the cell line MB66 (ref. 25) was screened under low stringency (hybridization in 30% formamide, 100 µg ml⁻¹ salmon sperm DNA, 5 × Denhardt's, 0.1% SDS, 20 mM Na phosphate, pH 6.8, and 1 M NaCl; washes in 150 mM NaCl, 15 mM Na citrate, all at 42 °C) with a nick-translated *v-jun* probe. Standard recombinant DNA techniques²⁶ were used to isolate and subclone the hybridizing cDNAs. All insert cDNAs were subcloned into the plasmid, pBI-76 and for nucleotide sequencing, a set of overlapping deletions were prepared by the Exo III-mung bean nuclease method of Henikoff²⁷. In certain instances, specific restriction fragments were subcloned and sequenced. Both strands of the cDNAs were sequenced. All nucleotide sequencing was done with the Sequenase kit (US Biochemicals) which utilizes the chain termination method of Sanger²⁸. Amino-acid-sequence alignments were carried out as described in ref. 24.



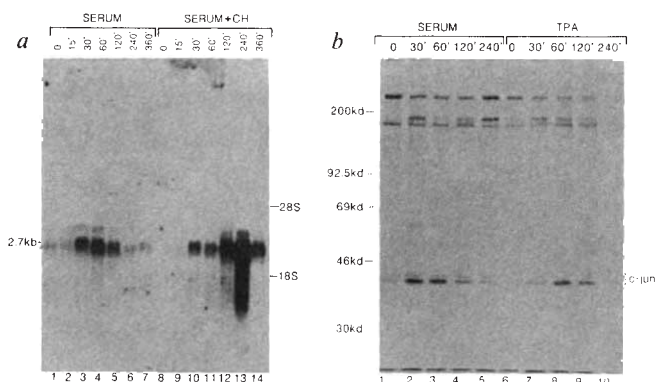


Fig. 2 Analysis of *c-jun* mRNA and protein following induction, *a*, RNA analysis. Autoradiogram of the kinetics of induction of *c-jun* mRNA with serum alone (lanes 1–7) and with serum plus cycloheximide ($10 \mu\text{g ml}^{-1}$) (lanes 8–14). *b*, Protein analysis. Immunoprecipitation of ³⁵S-labelled *c-jun* protein following addition of serum (lanes 1–5) or 12-O-tetradecanoyl-phorbol- β -acetate (TPA) (lanes 6–10).

Methods. Confluent 150 mm dishes of NIH 3T3 cells were serum-starved (0.5% serum) for 24–48 h prior to stimulation. For serum induction, the starved cells were induced with 20% fetal calf serum; for TPA induction, the starved cells were treated with 75 ng ml^{-1} TPA. Super-induced RNAs were isolated from cells treated with medium containing 20% fetal calf serum and $10 \mu\text{g ml}^{-1}$ cycloheximide. Total cellular RNA was isolated at the times indicated by the guanidine isothiocyanate method and Northern blot analysis of $10 \mu\text{g}$ (lanes 1–7) or $5 \mu\text{g}$ (lanes 8–14) of each RNA was carried out using 1% agarose formaldehyde gels²⁹. The RNA was transferred to nitrocellulose and hybridized with a 5' *c-jun* coding sequence probe (nucleotides 562–1,147, see Fig. 1) labelled with ³²P by nick-translation. The gel was exposed for three days. For protein analysis, confluent cell cultures were labelled with 1–2 mCi ml^{-1} of ³⁵S label (ICN Biochemicals). Prior to labelling, the cells were incubated for 20 min in medium lacking L-methionine and L-cysteine. The cells were then stimulated as described above, and at the times indicated, were pulse-labelled for 15 min. After labelling, nuclear lysates and immunoprecipitation was carried out as described in ref. 22. The gel was exposed for 16 h.

v-jun specific portion of ASV-17 (ref. 8). The cDNA codes for a protein that is 334 amino acids long, predicting a relative molecular mass (M_r) of ~37,000 (37K), which is in agreement with the observed molecular weight^{14,15}. The amino-acid sequence shares 98% identity with human c-JUN⁹ and 86% identity with v-JUN (Fig. 1b). It also shows an overall 50% amino-acid identity with the serum-inducible JUN B protein¹⁸. Compared to mouse or human *c-jun*, the *v-jun* gene has a number of deletions, the largest of which is a 27-amino-acid gap between positions 29 and 55 (Fig. 1b). Maximum sequence conservation is found in the carboxy-terminal domain of the JUN-related proteins, which is also the region showing similarity to the yeast transcriptional factor GCN4¹⁹. By analogy to GCN4, the carboxy-terminal domain of the murine JUN protein is likely to be the DNA binding domain. The 3'-untranslated region contains three poly(A) addition signal sequences (shown underlined in Fig. 1). All three poly(A) sites are functional, as of the four cDNA clones characterized, one used the first site, one used the second site, and the remaining two used the last site (data not shown). There may be still other poly(A) addition signal sequences in the *c-jun* mRNA. Many early response genes contain the sequence ATTTA, which has been implicated in mRNA instability in their 3' non-coding region²⁰. There is at least one in the murine *c-jun* cDNA clone (see open bar in Fig. 1).

Several studies have recently shown that *c-jun* codes for the *fos*-associated protein p39 (refs 14, 15), and as *c-fos* is an inducible gene, we conducted experiments to determine whether murine *c-jun* was also inducible. Figure 2 shows the kinetics of

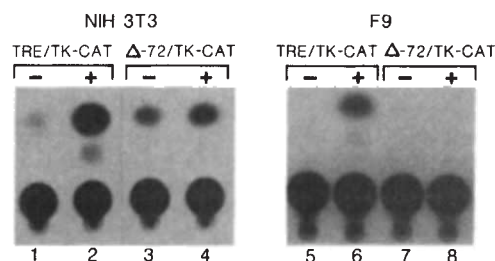


Fig. 3 The mouse *c-jun* gene codes for a functional AP-1 protein. The *c-jun* cDNA was cloned in a SV40-based expression vector²⁵. Clone pSV-*c-jun*, which contains a truncated *c-jun* cDNA (starting at position 611 in the cDNA sequence) was used in co-transfection assays. Plasmid TRE/TK-CAT contains a human metallothionein IIA TPA-responsive sequence element (the TRE consensus is GTGACTCAG) inserted upstream of the herpes thymidine kinase promoter (positions -109 to +57) which is fused to the CAT gene¹⁷. Co-transfection of pSV-*c-jun* (+), or expression vector only (-), with TRE/TK-CAT followed by CAT assays, determines the induction of transcription from the TK promoter in both NIH 3T3 (lanes 1–2) and F9 (lanes 5–6) cells. The use of a mutated TRE element (the TRE consensus is mutated to TGGAGTCAG in Δ-72/TK-CAT) demonstrates the specificity of the transcriptional activation (lanes 3–4 and 7–8).

Methods. Transfections were performed using the calcium-phosphate co-precipitation technique as described³⁰. Equal molar amounts of reporter plasmid and transactivator plasmid were transfected, reaching a total of $20 \mu\text{g}$ of DNA per 10-cm culture dish. CAT assays were performed as described³¹.

induction of *c-jun* mRNA by Northern blot analysis, following the addition of serum to growth-arrested NIH 3T3 cells (Fig. 2a, lanes 1–7). The steady state levels of *c-jun* RNA begin to increase at 30 min, reach a maximum at ~60 min, and then decrease, reaching pre-stimulation levels by 2 h. These kinetics are reminiscent of *c-fos* transcriptional induction except that in the case of *c-fos* there is no mRNA detectable in quiescent cells^{3–5}, whereas *c-jun* mRNA is easily detectable (see Fig. 2a, lane 1). Like *c-fos* mRNA, *c-jun* mRNA is super-inducible in the presence of $10 \mu\text{g ml}^{-1}$ cycloheximide (Fig. 2, lanes 8–14). Thus, *c-jun* mRNA is induced by agents that stimulate *c-fos* gene expression and the kinetics of induction of both genes are similar.

It has previously been reported that in FOS immunoprecipitates the level of labelled p39 (AP-1) protein increases in parallel to the level of the *fos* protein^{21,22}. The kinetics of the increase of p39 could not be directly investigated in the past due to the lack of specific reagents for p39. Figure 2b shows immunoprecipitation analysis, using anti-JUN peptide serum (anti-PEP1 and anti-PEP2, ref. 8), of nuclear lysates from NIH 3T3 cells stimulated with serum (lanes 1–5) or TPA (lanes 6–10). The synthesis of the 39K/c-JUN protein coincides with the induction of the mRNA (compare Fig. 2a and b). The c-JUN protein appears to be unstable with a half-life of about 90–120 min. The primary p39/c-JUN product appears to be post-translationally modified to a 42K protein. We also detect a 180–190K c-JUN-related protein which appears to be inducible, and post-translationally modified. This unidentified protein appears to be more stable than the *c-jun* protein.

To investigate whether the murine *c-jun* cDNA codes for a functional AP-1, able to transcriptionally activate genes containing TPA-responsive sequences (TRE), we performed co-transfection assays. A reporter plasmid TRE/TK-CAT (see legend to Fig. 3) was co-transfected with a *c-jun* expression vector (pSV-*c-jun*) in both NIH 3T3 cells and undifferentiated F9 embryonal carcinoma cells. As a control, plasmid containing mutations in the TRE motif of the TRE/TK-CAT plasmid was used (Δ-72/TK-CAT), ref. 17. As shown in Fig. 3, co-transfection with the *c-jun* expression vector leads to a significant increase

in CAT activity from the recombinant plasmid containing the functional AP-1 site (lanes 1-2 and 5-6) as compared to the one containing the mutated TRE (lanes 3-4 and 7-8). The extent of induction in the two cell lines is different because F9 cells have lower basal levels of AP-1 (ref. 23).

We have cloned the murine homologue of the viral *jun* oncogene (also referred to as PEA1, ref. 23), which codes for a functional AP-1 transcriptional factor. We suggest that murine *c-jun* is a member of the early response gene family by virtue of sharing the following common features: (1) its expression is inducible with growth factors and differentiation inducers; (2) the induction is very rapid and transient; (3) its mRNA contains the destabilizing sequence AUUUA in the 3'-non coding region which could contribute towards its short half-life; (4) the primary translational product is modified; and (5) the protein has a short half-life. Because of its inducibility by TPA (Fig. 2), and its ability to transcriptionally activate TPA-responsive genes, we can propose a model of positive transcriptional autoregulation of this gene. Understanding the response of the cell to external stimuli is fundamental in deciphering the mechanisms of growth control. It is likely that a highly organized network of nuclear proteins, some of which are products of oncogenes, governs the complex mechanism of gene expression. A serum, inducible mouse *c-jun* cDNA with a nucleotide sequence identical to the one reported here has also been identified (K. Ryder and D. Nathans, personal communication).

We thank H. Okayama for providing the cDNA expression library, T. Bos and P. Vogt for the generous gift of the *v-jun* probe and PEP-1 and PEP-2 antibodies, and to P. Angel, E. Allegretto and M. Karin for plasmids TRE/TK-CAT, Δ -72/TK-CAT. We thank J. Stewart for help in preparation of the manu-

script. W.W.L. is supported by the NIH; P.S.-C. is supported by the CNRS and is on leave from the Laboratoires de Génétique Moléculaire des Eucaryotes, CNRS, Strasbourg, France. This work was supported by grants from the NIH and American Cancer Society to I.M.V.

Received 16 June; accepted 1 July 1988.

1. Cochran, B. H., Reffel, A. C. & Stiles, C. D. *Cell* **33**, 939-947 (1983).
2. Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603-610 (1983).
3. Greenberg, M. E. & Ziff, E. B. *Nature* **311**, 433-438 (1984).
4. Kruijer, W., Cooper, J. A., Hunter, T. & Verma, I. M. *Nature* **312**, 711-716 (1984).
5. Muller, R., Bravo, R., Burckhardt, J. & Curran, T. *Nature* **312**, 716-720 (1984).
6. Lau, L. F. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1182-1186 (1987).
7. Almendral, J. M. *et al. Molec. cell. Biol.* **8**, 2140-2148 (1988).
8. Bohmann, D. *et al. Science* **238**, 1386-1392 (1987).
9. Angel, P. *et al. Nature* **332**, 166-171 (1988).
10. Sap, J. *et al. Nature* **324**, 635-640 (1986).
11. Weinberger, C. *et al. Nature* **324**, 641-646 (1986).
12. Rauscher, F. J. III, Sambucetti, L. C., Curran, T., Distel, R. J. & Spiegelman, B. M. *Cell* **52**, 471-480 (1988).
13. Franza, R. B., Jr., Rauscher, F. J. III, Josephs, S. F. & Curran, T. *Science* **239**, 1150-1153 (1988).
14. Rauscher, F. J. III *et al. Science* **240**, 1010-1016 (1988).
15. Sassone-Corsi, P., Lamph, W. W., Kamps, M. & Verma, I. M. *Cell* **54** (in the press).
16. Lee, W., Mitchell, P. & Tjian, R. *Cell* **49**, 741-752 (1987).
17. Angel, P. *et al. Cell* **49**, 729-739 (1987).
18. Ryder, K., Lau, L. F. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1487-1491 (1988).
19. Vogt, P. K., Bos, T. J. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3316-3319 (1987).
20. Shaw, G. & Kamen, R. *Cell* **46**, 659-667 (1986).
21. Curran, T., Miller, A. D., Zokas, L. & Verma, I. M. *Cell* **36**, 259-268 (1984).
22. Barber, J. R. & Verma, I. M. *Molec. cell. Biol.* **7**, 2201-2211 (1987).
23. Kryszke, M.-H., Piette, J. & Yaniv, M. *Nature* **328**, 254-256 (1987).
24. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**(1), 387-395 (1984).
25. Okayama, H. & Berg, P. *Molec. cell. Biol.* **3**, 280-289 (1983).
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, 1982).
27. Henikoff, S. *Gene* **28**, 351-359 (1984).
28. Sanger, F. *et al. Nature* **265**, 687-695 (1977).
29. Mitchell, R. L., Zokas, L., Schreiber, R. D. & Verma, I. M. *Cell* **40**, 209-217 (1985).
30. Sassone-Corsi, P. & Verma, I. M. *Nature* **326**, 507-510 (1987).
31. Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).

Yeast activators stimulate plant gene expression

Jun Ma*, Elisabeth Przbilla†, Jim Hu†, Lawrence Bogorad† & Mark Ptashne*

* Department of Biochemistry and Molecular Biology and
† Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138, USA

GAL4 is a transcriptional activator found in yeast^{1,2}. Two distinct functions of the protein are required for its activity³⁻⁵: one directs sequence-specific DNA binding, and another interacts with some other component of the transcriptional machinery, for example, RNA polymerase II or a TATA-binding protein. Two short regions of GAL4 function as 'activating sequences' when attached to the DNA-binding portion of GAL4 (ref. 6) and these regions can be replaced by a large number of peptides encoded by *Escherichia coli* genomic DNA fragments⁷ or by a synthetic peptide designed to form an amphiphilic α -helix⁸. All of these activating sequences, like that found in another yeast activator, GCN4 (refs 9, 10) bear an excess negative charge (also see ref. 11). GAL4 and its derivatives that are active in yeast stimulate transcription in mammalian cells when GAL4 binding sites are introduced upstream of a mammalian gene^{12,13}; similarly, GAL4 activates transcription in *Drosophila* cells¹⁴. Here we show that GAL4 derivatives stimulate gene expression in plant cells.

Figure 1 shows the three types of DNA constructs used in this study. The first type, the 'effector', expresses various GAL4 derivatives from the nopaline synthase (NOS) promoter¹⁵. These derivatives include: GAL4(1-147), the amino terminal portion of the 881-amino-acid GAL4 protein, which binds DNA but fails to activate transcription in yeast and in mammalian cells^{4,6,12}; GAL4(1-147)+II, a protein containing activating

region II (residues 768-881) fused to the DNA-binding portion⁶; GAL4(1-147)+I+II, a protein that bears both activating regions I (residues 148-196) and II fused to the DNA-binding portion⁶; and GAL4(1-147)+B17, which contains an activating sequence encoded by an *E. coli* DNA fragment fused to the DNA-binding portion of GAL4 (ref. 7). The second type of construct, the 'reporter', bears the transposon 9 chloramphenicol acetyl transferase (CAT) gene¹⁶ fused to the cauliflower mosaic virus (CMV) 35S promoter containing its TATA box and transcription initiation site¹⁷. One of the reporters (17mers-TATA-CAT) contains nine 17-base pair GAL4 binding sites (called 17mers; ref. 18) upstream of its TATA sequence, and the other (TATA-CAT) lacks these sites. The third type of construct, the internal control plasmid, contains the *E. coli* glucuronidase (GUS) gene¹⁹ expressed from the CMV 35S promoter with its normal enhancer²⁰.

We introduced three plasmids, one effector, one reporter, and the internal control plasmid, into tobacco-leaf protoplasts²¹ by an electroporation technique²². We assayed CAT activity in the extracts of the transfected plant protoplasts^{16,22}; the electroporation efficiency was normalized according to the GUS activity in each experiment²⁰.

Figure 2 shows that GAL4 derivatives bearing one or more activating sequences induced substantial CAT activity when GAL4 binding sites were located upstream of the plant gene (lanes 3-5). The DNA binding portion of GAL4 alone failed to induce any CAT activity (lane 2). We did not detect any CAT activity in the absence of GAL4 binding sites (lanes 7-11) or when no activator was expressed (lane 1). The CAT activity induced by the GAL4 binding sites in the presence of an active GAL4 derivative is comparable to that induced by a natural, and strong, plant enhancer—the CMV 35S enhancer (compare lanes 3, 4 and 5 with lane 6). Wild-type UAS_G (the GAL upstream activating sequence), which contains four natural binding sites¹⁸, also mediated gene activation (see legend to Fig. 2).

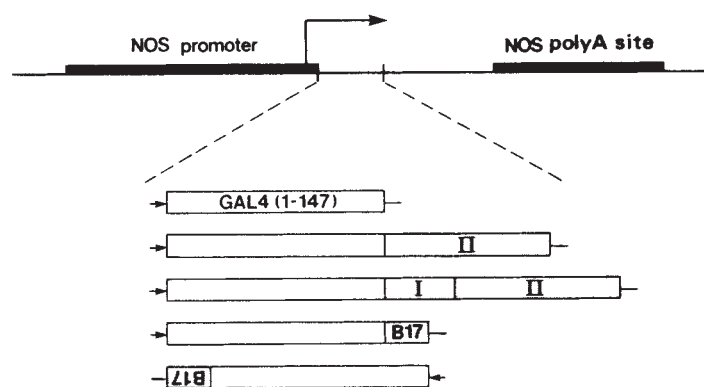
Our results extend the general conclusion that acidic activat-

Fig. 1 DNA constructs. *a*, 'Effectors'. Different GAL4 derivatives are expressed from the NOS promoter (refs 15, 22): GAL4(1-147), GAL4(1-147)+II, GAL4(1-147)+I+II and GAL4(1-147)+B17 are carried on pMA564, pMA562, pMA563 and pMA560, respectively. pMA560R contains the GAL4(1-147)+B17 sequences in the opposite orientation relative to the promoter, serving as a no-activator control. *b*, 'Reporters'. These plasmids bear the CAT gene¹⁶ fused to the CMV 35S promoter^{17,22} containing the TATA box and transcription initiation site. One of them, 17mers-TATA-CAT (pMA558), contains nine 17-base pair GAL4 binding sites inserted about 40 base pair upstream of the TATA sequence, and the other, TATA-CAT (pMA556), lacks the GAL4 binding sites. Not shown here is a 35S-CAT fusion (pCaMVCAT; ref. 22) that carries the normal enhancer of the CMV 35S gene. *c*, Internal control plasmid (pBI221; ref. 20). The GUS gene is expressed from the 35S promoter carrying its normal enhancer. The constructs are not drawn to scale.

Methods. *a*, 'Effectors' pMA560 and pMA560R were constructed by inserting *Hind*III-*Sal*I fragment of plasmid B17 (ref. 7) into the *Bam*HI-*Nco*I site of pNOSCAT (ref. 22), a plasmid bearing the sequences of NOS promoter-CAT-NOS polyadenylation site (poly(A) site); the restriction ends had been filled in with Klenow. pMA560 expresses GAL4(1-147)+B17 from the NOS promoter, whereas pMA560R contains the GAL4(1-147)+B17 sequences in the opposite orientation. pMA562, pMA563 and pMA564 were constructed by replacing the *Xho*I-*Sal*I sequence of pMA560 with the *Xho*I-*Hind*III fragments of pMA236, pMA242 and pMA241 respectively⁶; *Sal*I and *Hind*III ends had been filled in with Klenow. *b*, 'Reporters'. pMA556 and pMA558 were constructed as follows. The *Eco*RV-*Hind*III fragment of pCaMVCAT²² containing the sequences of the 35S promoter-CAT-NOS poly(A) site was inserted at the *Hinc*II-*Hind*III sites of pUC18, generating pMA554. pMA556 was constructed by deleting the *Xba*I-*Hga*I fragment, which contains the putative CAAT box¹⁷ of pMA554. pMA558 was constructed by inserting 17-base pair oligonucleotides of GAL4 binding sites (called 17mers), kindly provided by A. Miller and M. Carey, at the *Sma*I site of pMA556; nine to ten 17-base pair GAL4 binding sites were inserted, as estimated by restriction mappings.

a Effectors

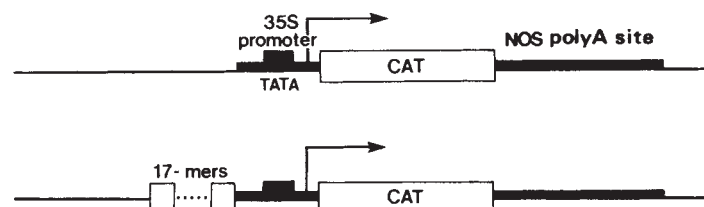
GAL4 (1-147)
GAL4 (1-147)+II
GAL4 (1-147)+I+II
GAL4 (1-147)+B17
no activator



b Reporters

TATA-CAT

17mers-TATA-CAT

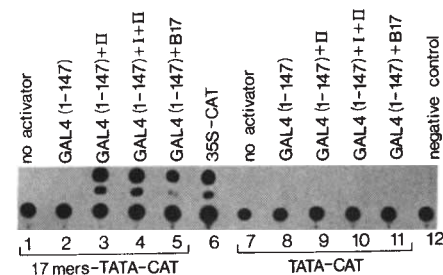


c Internal control plasmid



Fig. 2 CAT assay. Three plasmids, one effector, one reporter and the internal control plasmid, were introduced into tobacco-leaf protoplasts²¹ by electroporation technique²². One of the reporter plasmids contained the 17-base pair GAL4 binding sites (17mers-TATA-CAT; lanes 1-5); the other did not (TATA-CAT; lanes 7-11). The effectors included: no-activator controls (lanes 1 and 7), GAL4(1-147) (lanes 2 and 8), GAL4(1-147)+II (lanes 3 and 9), GAL4(1-147)+I+II (lanes 4 and 10) and GAL4(1-147)+B17 (lanes 5 and 11). Lane 6 shows the result obtained from protoplasts transfected with the intact 35S-CAT fusion (pCaMVCAT; ref. 22) bearing the normal 35S enhancer. Lane 12 shows a negative control that was not transfected with either a reporter or an effector. All the experiments demonstrating the activation by our GAL4 derivatives were performed at least twice with similar results; the activation by GAL4(1-147)+II was repeated five times with similar results. For wild-type GAL4, we observed very little, if any, stimulation of CAT activity in our experiments (data not shown); we do not know whether this is caused by, for example, inefficient translation or protein instability. For all the experiments in Fig. 2 the putative CAAT box located upstream of the TATA box was deleted. In the presence of this CAAT box there was a detectable basal activity in the absence of GAL4-derived activators. Although GAL4(1-147)+II activated transcription, the final level was somewhat lower than those achieved in the experiments shown here. We also reproducibly detected activation by GAL4(1-147)+II on a reporter that contained the wild-type UAS_G located upstream of the CAAT box, and the activation was at least as significant as on the reporter containing the synthetic 17 base pair repeats upstream of the CAAT box (data not shown).

Methods. Tobacco protoplasts were isolated from leaves as previously described²¹. For each electroporation an electric pulse of 350 V cm⁻¹ and 25 ms was delivered by a 1,000 µF capacitor-discharge circuit (Promega 450) to about 10⁶ protoplasts in Hepes-buffered saline²² with 40 µg of effector, 15 µg of reporter, and 5 µg of the internal control plasmid (pBI221; ref. 20). The culture was incubated at 25 °C in the dark for about 40 hours and the lysates were tested for both GUS and CAT activities^{16,20,22}; CAT assay was normalized by GUS activity in each experiment.



ing sequences function in many eukaryotic systems and reinforce the likelihood of a common mechanism of transcriptional activation in eukaryotes. They also suggest that plant activators might function (given an appropriate DNA-binding site) in yeast, which would provide a useful tool for studying plant regulatory genes.

We thank members of the Ptashne and Bogorad labs for

helpful discussions; Allan Miller and Mike Carey for the oligonucleotides of the GAL4 binding sites; Richard Jefferson for the internal control plasmid; Betsy Burkhardt for artwork; and Tony Kavanagh for communicating unpublished results. This work has been supported by grants from NIH to M.P. and NIH to L.B. E.P. was supported by a grant of the special programme Gentecnologie of the DAAD.

Received 3 May; accepted 27 June 1988.

- Oshima, Y. in *Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression* (eds. Strathern, J. N., Jones, E. W. & Broach, J. R.) 159-180 (Cold Spring Harbor Laboratory, New York, 1982).
- Johnston, M. *Microbiol. Rev.* **51**, 458-476 (1987).
- Brent, R. & Ptashne, M. *Cell* **43**, 729-736 (1985).
- Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
- Ptashne, M. *Nature* **322**, 697-701 (1986).
- Ma, J. & Ptashne, M. *Cell* **48**, 847-853 (1987).
- Ma, J. & Ptashne, M. *Cell* **51**, 113-119 (1987).
- Giniger, E. & Ptashne, M. *Nature* **330**, 670-672 (1987).
- Hope, I. & Struhl, K. *Cell* **46**, 885-894 (1986).
- Struhl, K. *Cell* **49**, 294-297 (1987).
- Gill, G. & Ptashne, M. *Cell* **51**, 121-126 (1987).
- Kakidani, H. & Ptashne, M. *Cell* **52**, 161-167 (1988).
- Webster, N., Jin, J. R., Green, S., Hollis, M. & Chambon, P. *Cell* **52**, 169-178 (1988).
- Fisher, J., Giniger, E., Maniatis, T. & Ptashne, M. *Nature* **332**, 853-856 (1988).
- Bevan, M., Barnes, W. M. & Chilton, M. D. *Nucleic Acids Res.* **11**, 369-385 (1983).
- Alton, N. K. & Vapnek, D. *Nature* **282**, 864-869 (1979).
- Odell, J. T., Nagy, F. & Chua, N.-H. *Nature* **313**, 810-812 (1985).
- Giniger, E., Varnum, S. & Ptashne, M. *Cell* **40**, 767-774 (1985).
- Jefferson, R. A., Burgess, S. M. & Hirsch, D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8447-8451 (1986).
- Jefferson, R. A. *Pl. molec. Biol. Reporter* **5**, 387-405 (1987).
- Nagy, J. I. & Maliga, P. Z. *Pflanzenphysiol.* **78**, 435-455 (1976).
- Fromm, M., Taylor, L. P. & Walbot, V. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5824-5828 (1985).

Long-range molecular order as an efficient strategy for light harvesting in photosynthesis

Zoya G. Fetisova*, Arvi M. Freiberg†
& Kiu E. Timpmann†

* A. N. Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 119899, USSR

† Institute of Physics, Estonian SSR Academy of Sciences, Tartu 202400, USSR

Photosynthesis is an extremely efficient converter of light into chemical energy, with an observed quantum yield for primary photochemistry in the region of 90%. To achieve this the photosynthetic apparatus must be highly optimized¹, and some of the design principles that may be involved have been suggested². A search for the realization of these principles in natural photosynthetic units led to the discovery of predicted strong orientational ordering of the long-wave (Q_y) transition moments of light-harvesting bacteriochlorophyll *c* in the green bacterium *Chlorobium limicola*^{3,4}. These transition moments have been shown by linear dichroism to be parallel to each other and ideally oriented along the long axis of the chlorosome. A conclusion of such importance needs to be tested in an independent experiment with intact cells, and this we report here. We have directly measured the picosecond polarized fluorescence kinetics for bacteriochlorophyll emissions upon selective polarized light excitation of the Q_y transition of bacteriochlorophyll *c* fluorescence polarization value was found to be *Chlorobiaceae* and *Chloroflexaceae*. The measured bacteriochlorophyll *c* fluorescence polarization value was found to be constant during the lifetime of the bacteriochlorophyll *c* excited state, reaching a limiting value for monomeric bacteriochlorophyll of 0.42. These results demonstrate that in living cells of green bacteria excitation energy transfer within the bacteriochlorophyll *c* antenna occurs between chromophores (or their associates) with parallel transition moments.

The green sulphur bacteria were chosen for a linear-dichroism study^{3,4}, as, in these bacteria, the light-harvesting antennae are an order of magnitude larger than in purple bacteria, and several-fold larger than in higher plants. An investigation of the ordering in these large and efficient photosynthetic units (PSUs) is therefore of value, as the requirements for structural optimization are more rigorous than for small ones¹. The green bacteria are classified in two families, *Chlorobiaceae* and *Chloroflexaceae*, which differ in terms of both evolution and physiology⁵. *Chloroflexaceae* is thought to be a family of transitional organisms

between purple bacteria, green sulphur bacteria, and cyanobacteria^{5,6}. In the work presented here we have studied one species from each family of green bacteria: the sulphur bacterium *Chlorobium limicola* and the thermophilic filamentous non-sulphur bacterium *Chloroflexus aurantiacus*. The two species differ in the structural organization of that part of the PSU which is localized in the cytoplasmic membrane, but are similar in their extra membrane antenna structures: ellipsoid bodies attached to the inside of the cytoplasmic membrane called chlorosomes⁵⁻⁹. Chlorosomes contain the major antenna BChl *c*, whereas the reaction centres and the greater part of the antenna BChl *a* are associated with the cytoplasmic membrane^{6,7}. Excitation energy transfer from the chlorosome pigments to those of the membrane is very rapid and highly efficient in both species¹⁰⁻¹³.

To examine the long-range orientational ordering of Q_y transition moments of BChl *c* throughout the entire chlorosome, we used picosecond polarized fluorescence spectroscopy¹⁴. This was the most informative and suitable technique for our purposes because it could be applied to photosynthetic material *in situ*.

All the experiments were performed at room temperature on intact 2-3-day-old cells in their own growth medium under strictly anaerobic conditions. The absorption and fluorescence spectra of both *Chlorobium* and *Chloroflexus* species are shown in Fig. 1. The picosecond polarized fluorescence kinetics for antenna BChl emissions in living cells was measured upon selective excitation with polarized light of the Q_y transition of BChl *c*. The fluorescence polarization function, $p(t)$, was calculated according to the equation $p(t) = [I_{\parallel}(t) - I_{\perp}(t)] / [I_{\parallel}(t) + I_{\perp}(t)]$, where $I_{\parallel}(t)$ and $I_{\perp}(t)$ are the signal intensities associated with respective fluorescence components measured through a polarizer oriented either parallel or perpendicular to the polarization direction of the exciting light. In *C. limicola* cells the excitation at 711 nm was selected, and the picosecond BChl *c* and BChl *a* fluorescence kinetics were detected at 730 and 820 nm, respectively. In *C. aurantiacus* cells the excitation at 724 nm was selected, and the picosecond BChl *c* fluorescence kinetics were detected at 750 nm; the BChl *a* fluorescence kinetics were detected at both BChl *a* fluorescence maxima wavelengths, at 806 and 880 nm.

Figures 2 and 3 show the picosecond polarized fluorescence kinetics for BChl *c* and BChl *a* emissions in living cells of *C. limicola* and *C. aurantiacus*, respectively, upon selective excitation of the BChl *c* Q_y transition, as well as the corresponding calculated fluorescence polarization function, $p(t)$. In *C. limicola* cells, the BChl *c* fluorescence polarization function is constant during the lifetime of the excited state, with an average value of 0.422 ± 0.003 in the time interval 0-280 ps.

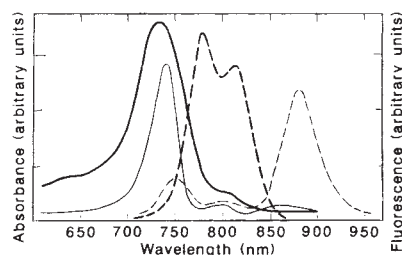


Fig. 1 Room-temperature near-infrared absorption (solid lines) and uncorrected fluorescence (dashed lines) spectra of intact cells of *C. limicola*, strain C (thick lines), and of *C. aurantiacus*, strain B III (thin lines) under strictly anaerobic conditions. Excitation was at 460 nm. For *C. limicola*: the absorption maximum at 733 nm and the fluorescence maximum at 780 nm belong to BChl *c*; the absorption maximum at 805 nm and the fluorescence maximum at 812 nm belong to BChl *a*. For *C. aurantiacus*: the absorption maximum at 740 nm and the fluorescence maximum at 750 nm belong to BChl *c*; the absorption maxima at 798 and 867 nm and the fluorescence maxima at 803 and 880 nm belong to BChl *a*.

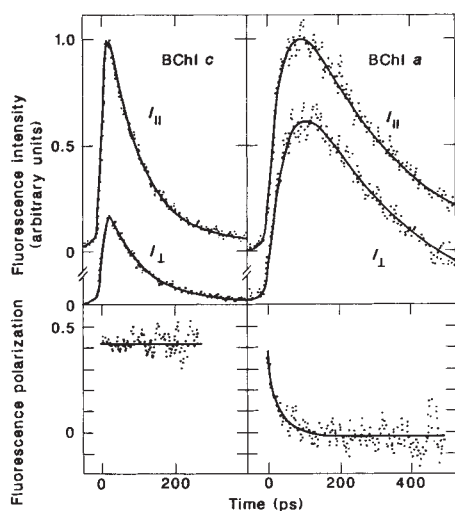


Fig. 2 Polarized fluorescence kinetics ($I_{||}$ and $I_{\perp}$, dotted curves, vertically shifted for clarity) for BChl *c* (at 730 nm) and BChl *a* (at 820 nm) fluorescence decay in living cells of *C. limicola*. Excitation was at 711 nm. The solid curves are the best multi-exponential fits to the data (within an accuracy of 15%). Both BChl *c* kinetics, $I_{||}$ and $I_{\perp}$, are strongly biphasic and characterized by a fast phase with a lifetime $\tau_f = 90$ ps and a slow phase with a lifetime $\tau_s = 300$ ps (the amplitude ratio $I_f/I_s = 6$). Both BChl *a* decay kinetics are approximated by a single exponent, with $\tau = 300$ ps. Bottom, the corresponding calculated fluorescence polarization functions, $p(t)$, where $p = (I_{||} - I_{\perp}) / (I_{||} + I_{\perp})$.

Methods. The picosecond spectrochronograph used has been described¹⁴. Briefly, the picosecond pulse source was a mode-locked CW oxazine 1 dye laser (pulse duration, 3 ps), synchronously pumped at 82 MHz by a krypton-ion laser. The excitation pulses were vertically polarized. Emission, viewed at an angle of 90° to the exciting beam in a reflection mode, was filtered by two single-grating monochromators (bandwidth of 4 nm) and was registered by a synchroscan streak camera. The time resolution of the apparatus was 3–5 ps. For polarization measurements two polarizers were used. The excitation intensity was equal to 0.2 W cm^{-2} . The kinetics observed were unchanged over an approximately 10^3 -fold decrease in the laser pulse intensity, and so no exciton annihilation occurred.

For *C. aurantiacus* cells the main finding is the same: the mean value $p(t)$ was 0.41 ± 0.01 in the time interval 0–230 ps. The insignificant decrease in the polarization function in this time interval from 0.45 to 0.41 is due to a relatively large contribution to the measured biphasic decay kinetics of the slow component ($I_f/I_s = 1.5$ compared to 6.0 in *Chlorobium*) existing because of the back excitation transfer. The fairly high degree of BChl *c* fluorescence polarization ($p = 0.33$ – 0.37) was obtained earlier in steady-state fluorescence measurements performed at 4 K on isolated chlorosomes of *C. aurantiacus*¹⁵.

Owing to the significant overlapping of BChl *c* and BChl *a* emission bands at 820 nm (for *C. limicola*) and at 806 nm (for *C. aurantiacus*), the $p(t)$ decay kinetics at these wavelengths reflect the kinetics of the change in the ratio of BChl *c* and BChl *a* emission intensities during the first 100 ps. For $t > 100$ ps, $p(t)$ approaches 0 for both species. For *C. aurantiacus* cells, the BChl *a* fluorescence polarization function at 880 nm (not shown) was constant throughout the excited-state decay time ($\tau_d = 350$ ps) and equalled 0.06 ± 0.03 .

The constant value of $p(t)$ obtained for BChl *c* emission reached a limiting value of p for monomeric BChl molecules¹⁶. This result suggests that in living cells of green bacteria the excitation energy transfer within the BChl *c* antenna (contained in the chlorosome) occurs between chromophores (or their strongly coupled associates) with parallel transition moments. It is likely that these transition moments are those of the collective excitation within a cluster of several strongly coupled BChl

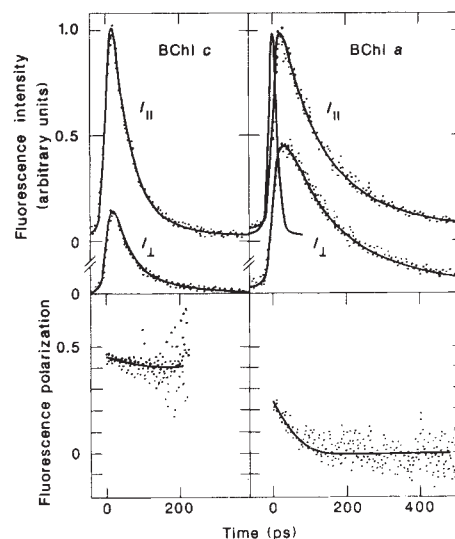


Fig. 3 Polarized fluorescence kinetics for BChl *c* (at 750 nm) and BChl *a* (at 806 nm) fluorescence decay in living cells of *C. aurantiacus*. Excitation was at 724 nm. Both BChl *c* kinetics are two-exponential, with $\tau_f = 30$ ps and $\tau_s = 100$ ps ($I_f/I_s = 1.5$). Both BChl *a* kinetics are also biphasic, with $\tau_f = 100$ ps and $\tau_s = 350$ ps ($I_f/I_s = 2.3$). The narrow profile (FWHM = 16 ps) is the apparatus response function. Bottom, the corresponding functions, $p(t)$.

c molecules¹⁷ rather than those of the individual chromophores. In this case the excitation energy transfer within the BChl *c* antenna may be described as that between those clusters with parallel transition moments: each cluster may be considered as a single large 'molecule' which (as a whole) may serve as a donor or acceptor molecule in hopping-type excitation transfer (as recognized for the light-harvesting chlorophyll *a/b* protein of higher plants¹⁸). It is this model that was used in computations of optimal orientation of BChl *c* transition moments in PSU of *C. limicola*⁴.

Our result is the first direct experimental demonstration of a high long-range molecular order throughout the whole light-harvesting structure in living cells. (Preliminary data on *C. limicola* were recently published¹⁹). The biological expedience of such an order in natural PSUs is clear, since only ordered systems may be optimized. Our result implies that in natural PSUs, where the pigment antenna 'lattice' is governed by the three-dimensional structure of the pigment-bearing proteins, the antenna proteins serve to hold the pigments in an exquisitely precise mutual position ensuring directionality of the energy transfer toward the reaction centres.

Received 29 February; accepted 22 June 1988.

1. Fetisova, Z. G. & Fok, M. V. *Molec. Biol.* (Engl. transl.) **18**, 1354–1359 (1984).
2. Fetisova, Z. G., Borisov, A. Yu. & Fok, M. V. *J. theor. Biol.* **112**, 41–75 (1985).
3. Fetisova, Z. G., Kharchenko, S. G. & Abdourakhmanov, I. A. *FEBS Lett.* **199**, 234–236 (1986).
4. Fetisova, Z. G., Kharchenko, S. G. & Abdourakhmanov, I. A. in *Progress in Photosynthesis Research* Vol. 1 (ed. Biggins, J.) 415–418 (Nijhoff, Dordrecht, 1987).
5. Fox, G. E. *et al. Science* **109**, 457–463 (1980).
6. Feick, R. G., Fitzpatrick, M. & Fuller, R. C. *J. Bact.* **150**, 905–915 (1982).
7. Schmidt, K. *Arch. Microbiol.* **124**, 21–31 (1980).
8. Staehelin, L. A., Golecki, J. R., Fuller, R. C. & Drews, G. *Arch. Microbiol.* **119**, 269–277 (1978).
9. Staehelin, L. A., Golecki, J. R. & Drews, G. *Biochim. biophys. Acta* **589**, 30–45 (1980).
10. Fetisova, Z. G. & Borisov, A. Yu. *FEBS Lett.* **114**, 323–326 (1980).
11. Van Dorssen, R. J., Gerola, P. D., Olson, J. M. & Ames, J. *Biochim. biophys. Acta* **848**, 77–82 (1986).
12. Brune, D. C., King, G. H., Infosino, A., Steiner, T., Thewalt, M. L. W. & Blankenship, R. E. *Biochem. J.* **26**, 8652–8658 (1987).
13. Van Dorssen, R. J. & Ames, J. *Photosynth. Res.* **15**, 177–189 (1988).
14. Freiberg, A. M. *Laser Chem.* **6**, 233–252 (1986).
15. Van Dorssen, R. J., Vasmel, H. & Ames, J. *Photosynth. Res.* **9**, 33–45 (1986).
16. Breton, J. & Vermeglio, A. in *Photosynthesis: Energy Conversion by Plants and Bacteria* Vol. 1 (ed. Govindjee) 153–194 (Academic, New York, 1982).
17. Smith, K. M., Kehres, L. A. & Fajer, J. *J. Am. chem. Soc.* **105**, 1387–1389 (1983).
18. Knox, R. S. & Van Metter, R. L. in *Chlorophyll Organization and Energy Transfer in Photosynthesis* (eds Wolstenholme, G. & Fitzsimons, D. W.) 177–190 (Excerpta Medica, Amsterdam, 1979).
19. Fetisova, Z. G., Freiberg, A. M. & Timpmann, K. E. *FEBS Lett.* **223**, 161–164 (1987).

Automated chromosome analysis

J.P. Zeidler

The increasing interest in and demand for the study of human genetics has led to the introduction on the market of a diverse array of systems for automated metaphase spread location and karyotyping.

THE defect of a single gene may cause fatal disease, so it is not surprising that the aberration or loss of even a part of a somatic chromosome precludes survival, except in very rare cases. Abnormal cell groupings within the body, particularly leukaemic bone marrow and malignant tumours, may show considerable and characteristic disturbances in their chromosome patterns. The examination and classification (karyotyping) of chromosomes from fetal cells is an essential preliminary to the detection and understanding of congenital disease arising from chromosomal defects. The microscopic examination of chromosomes in the metaphase of cell division is the method widely and routinely used for the detection of chromosomal aberrations in medical genetics, toxicological studies of mutagenicity, the evaluation of drug safety and radiation health monitoring.

The tasks involved in locating a number of individual metaphase spreads on a microscope slide, selecting and evaluating the most appropriate spreads, karyotyping and, when required, establishing a karyogram, are very time-consuming. In both routine and research environments the process can become tedious and tiring, however, the importance and consequences of the results of the screening process demand consistency and accuracy from highly trained and qualified staff.

The increasing demands on health services to provide genetic counselling and pregnancy screening, and the burgeoning requirements for mutagenicity testing and radiation health monitoring have added to the pressures on cytogenetic laboratory services. The combined effect of these factors, coupled with the omnipresent need to contain costs, has prompted and accelerated the design and development of instrument systems for automated metaphase spread location and karyotyping.

Early development

The first instrument for automated chromosome analysis was described in a paper published in 1963 by Robert Ledley¹. The instrument consisted of a film digitizer coupled to a computer equipped with software for the extraction of boundaries and the analysis of curvature sequences. The system created considerable interest, and prompted the Medical Research Council (MRC) in the United Kingdom to found a pattern recognition group in its Clinical and Population Cytogenetic Unit

in Edinburgh.

The MRC group was initially concerned with designing an instrument to locate metaphases. The first system was based on a large, custom-built "microscope" with a scanning stage using video imaging, which even early on achieved high search speeds. However, this was at the cost of a large amount of custom hardware, with a resultant lack of flexibility in performance. A system was subsequently developed in conjunction with a team at Tufts New England Medical Center, then led by Peter Neurath. This system was based on a more conventional optical microscope, and used the then new technology of a CCD array scanner².

Growth in the field of cytogenetics prompted several companies which were active in the field of video image analysis to produce adaptations of their existing universal instrument systems by developing specialized software to carry out metaphase spread location and karyotyping, with varying degrees of user interaction. These included Kontron GmbH in West Germany with the IBAS Video Image Analysis System, Cambridge Instruments in the United Kingdom with the Quantimet 720, and Ernst Leitz Wetzlar GmbH in West Germany, which introduced a metaphase spread and karyotyping program named METFIN in 1980 for their TAS+ System³. Around the same time, Joyce-Loebl in the United Kingdom developed an effective system for both metaphase location and karyotyping, based on the MAGISCAN Image Analysis System, in collaboration with John Phillip of the Rigshospitalet in Copenhagen, Denmark, and the Wolfson Image Analysis Unit in Manchester, United Kingdom⁴. A number of companies in the United States have also been active in this field and have enjoyed notable commercial success, including Perceptive Systems, Inc. in Houston, Texas and the Applied Imaging Corporation in Santa Clara, California.

Work continued to progress with the MRC group⁵⁻⁷ in Edinburgh in the design and construction of various prototype systems, and in early 1981 a company called Shandon Southern Products Ltd in the United Kingdom commissioned the building of two instrument systems (now marketed by Image Recognition Systems Ltd, Warrington, United Kingdom) which were given the name Cytoscan 110⁸.

The availability and relative reduction in the cost of computer memory and

framestores has facilitated the development of automated karyotyping, which is available, with a greater or lesser degree of user interaction, from the suppliers of the systems previously described. In addition, systems dedicated to the task of karyotyping have been developed — notably one named Karyotec, by the Israeli company Amcor — the first one of which has been installed in the United Kingdom at the Royal Masonic Hospital in London.

On the horizon

Progressive advances in computer hardware will undoubtedly continue to provide increases in the speed of operation of automated metaphase spread location and karyotyping systems. Today, most instruments have average scanning speeds of under 3-4 minutes per slide. In continuous 24-hour operation a typical instrument can scan a volume of slides that would take an operator using a conventional microscope up to one month to complete, depending on the material and number of karyotypes required.

With such instrument systems now starting to fulfill a central function in cytogenetic laboratories, the demand is emerging for "satellite" microscope systems to enable the cytogeneticist to obtain and interact with the findings of the machine without interrupting or disturbing its routine scanning functions. This requirement has not yet generally been met, but demand for such systems will ensure that they become readily available in the future.

The considerable pressures which led to the development of automated metaphase spread location and karyotyping systems will continue to increase. These, coupled with the increased speed of operation, reliability and convenience of use of the instruments, and their reduction in cost will ensure the deployment and use of such systems on an increasingly wide scale in cytogenetics laboratories. □

John P. Zeidler is at E. Leitz Instruments Ltd, 48 Park Street, Luton LU1 3HP, UK. For more information, fill in reader service number 100.

1. Ledley, R.S., Ruddle, F. & Migeon, B. in *Proc. 16th Ann. Conf. Eng. med. Biol.*, (PTBGME, New York, 1963).
2. Green, D.K. & Neurath, P.W. *J. Histochem. Cytochem.* **22**, 531 (1974).
3. Reuter, W., Stoll, D. *Labor-Praxis in der Medizin* **3**, (1980).
4. Foster, J. *Analyt. quant. Cytol.* **7**, 192 (1985).
5. Piper, J. & Rutovitz, D. *Pattern Recog. Lett.* **4**, 397 (1986).
6. Piper, J. & Rutovitz, D. *Pattern Recog. Lett.* **3**, 119 (1985).
7. Finnon, P., Lloyd, D.C. & Edwards, A. *Mutation Res.* **164**, 101 (1986).
8. Loerch, T., thesis, Heidelberg Univ. (1986).

World genetics conference in Toronto

Pulsed field gel electrophoresis, cell lines and automated chromosome sorters will be the talk of the town next week in Toronto, during the 16th International Congress of Genetics.

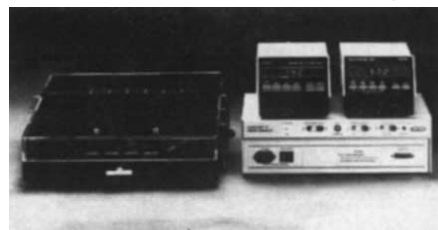
IN booth F-16, the American Type Culture Collection will have information on its latest **catalogue on cell lines and hybridomas** (Reader Service No. 101). The newly released sixth edition contains listings for over 2,800 characterized cell lines from approximately 75 species, including 150 human skin fibroblast lines from normal individuals and from patients with various genetic disorders. The catalogue also includes media formulations for growing each cell line. The catalogues are free to US researchers, but there is a \$15 (US) shipping and handling charge for foreign orders.

Becton Dickinson, in booths G-07 and G-09, has a **quantitative image analyser** which pairs software packages for routine clinical DNA ploidy and estrogen receptor measurements with versatile research software (Reader Service No. 102). The

\$34,000 (US) CAS 100 (not including software) is built around a Reichert Diastar microscope equipped with a CCD camera, and linked to a standard IBM PC/AT-3 with colour graphics presentation. Becton Dickinson says the system extracts highly reproducible cell measurements from images, to analyse cells and generate data about them in the form of histograms and scattergrams. The system's programs also perform statistical analyses to discriminate between subpopulations of cells, and all images can be stored for future reference. The CAS 100 comes complete with a reagent kit containing calibration slides, stains, single batch aliquot reagents and instructions.

Applied Biosystems now offers a **reagent kit for nucleic acid extraction** (Reader Service No. 103). The kit was first developed for use with the company's model 340A extractor, but it also may be used for manual extractions. Applied Biosystems says the kit eliminates the need to

screening (Reader Service No. 104). Its name stands for Dynamically Regulated Clamped Homologous Electric Field, which Bio-Rad says combines electronic and electrophoresis technologies to produce straight lanes and a size separation range that can resolve 15 yeast



The CHEF-DR II electrophoresis system from Bio-Rad can resolve 15 yeast chromosomes.

chromosomes. The CHEF-DR II possesses a hexagonal array of 24 electrodes, and each electrode is "clamped" to the desired level regardless of fluctuations in electrophoresis conditions, to yield up to 30 straight lanes in one run. Bio-Rad says the CHEF-DR II works best in the 2,000–1,000,000 base range, but that fragments as small as 222 bases have been separated. Bio-Rad is in booth F-03.

The Canadian company Cedarlane Laboratories, exhibiting in booths G-16 and G-18, has a new catalogue of **immunology research and diagnostic reagents** (Reader Service No. 105). Cedarlane is best known for its rabbit and guinea pig complement, sold under the trade name Low-Tox, but the company also offers cytotoxicity medium, lympholyte-M cell separation medium, antibodies to mouse T-cell surface antigens, antibodies to mouse MHC class I and class II antigens, and monoclonal antibodies to albumin, transferrin, and Tamm-Horsfall protein.

In booths B-01, B-02 and B-03, Image Recognition Systems will have information on its Cytoscan RK1 **rapid karyotyping analysis system** (Reader Service No. 106). The Cytoscan RK1 uses conventional video technology to digitize and process metaphase spread images. The company says it produces hard copies of karyograms using a digital video laser printer which are equivalent to or better than those obtained by conventional photographic means. The system can be used with any conventional microscope found in a standard cytogenetics laboratory. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

Economical Modular Tissue Culture Incubator

- Safe for AIDS research • Reliable, for In-Vitro fertilization
- Versatile enough for thousands of lab uses

Easy to use – Just flush with gas mixture, seal and place at approp. temp. Perform multiple experiments simultaneously w/out cross contamination. Thousands in use worldwide.

For information contact: **Flow Laboratories, Ltd.** or, **Billups-Rothenberg, Inc.** • P.O. Box 977 Del Mar, CA 92014-0977 USA • (619) 755-3309

Reader Service No.25

LIPOSOMES

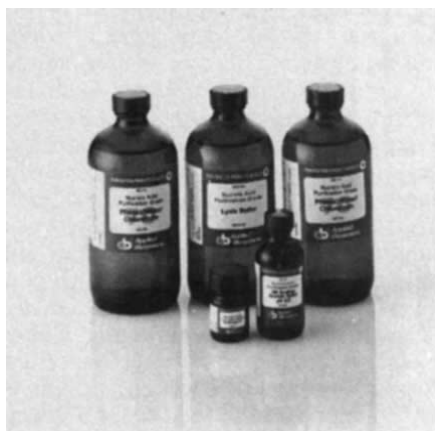
REPRODUCIBLE – HOMOGENEOUS – EXTREMELY RAPID

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.48



DNA extraction from Applied Biosystems.

prepare lysis solutions or to distill phenol with buffer, minimizing technician exposure to toxic substances. The kit includes a specially formulated lysis buffer, proteinase K, 70 per cent phenol/water/ chloroform solution, and 3 M sodium acetate buffer for "salting out" DNA from aqueous solutions. All of the reagents are provided pre-mixed, along with detailed protocols for the isolation of DNA from leukocytes, cell cultures and soft tissues. Applied Biosystems is in booths H-14, H-16 and H-18.

The CHEF-DR II **megabase DNA electrophoresis system** from Bio-Rad may be used in a variety of molecular biology applications, including the study of chromosome arrangements, restriction fragment length polymorphism analysis, insert sizing, gene mapping and library